DOCUMENT RESUME

ED 097 955 PS 007 163

AUTHOR Kirchner, Elizabeth P.
TITLE An Assessment Inventory for the Day Care Child. Volume II: Field Evaluation and Preliminary Findings. Center for Human Services Development, Report No. 25.
INSTITUTION Pennsylvania State Univ., University Park. Coll. of Human Development.
SPONS AGENCY Pennsylvania State Dept. of Public Welfare, Harrisburg.
REPORT NO PSU-CHSD-R-25
PUB DATE Jun 73
NOTE 199p.; For other reports in this series, see PS 007 154 - PS 007 164

EDRS PRICE MF-$0.75 HC-$9.00 PLUS POSTAGE
DESCRIPTORS Age Differences; *Behavior Rating Scales; Creative Development; *Day Care Programs; *Evaluation; Language Development; *Measurement Techniques; Motor Development; Observation; *Preschool Tests; Problem Solving; Racial Differences; Self Concept; Sex Differences; Social Development; Visual Discrimination

ABSTRACT

This is the second of a two volume report of a Day Care Inventory. The author presents an evaluation of the Inventory and make recommendations for changes, deletions and additions so that more satisfactory measures will be available for use in the study of the impact of day care on child development. The evaluations are based upon the findings from 282 day care children in Pennsylvania. The Day Care Inventory contains sixteen separate assessment components; these measures can be used to assess the development of preschool children in the following areas: (1) problem solving; (2) creativity; (3) sensory experience; (4) language; (5) gross motor skills; (6) visual discrimination; (7) social development and skills; (8) self esteem; (9) self concept; and (10) affective behavior. The format for each component evaluation includes a test description, discussion of the administration and scoring, a report of the findings, and a summary containing recommendations. (SDH)

INSTITUTE FOR THE STUDY OF HUMAN DEVELOPMENT

CENTER FOR HUMAN SERVICES DEVELOPMENT

AN ASSESSMENT INVENTORY FOR THE DAY CARE CHILD
VOLUME II -- FIELD EVALUATION AND PRELIMINARY FINDINGS

By: Elizabeth P. Kirchner

June, 1973

CHSD Report No. 25

COLLEGE OF HUMAN DEVELOPMENT
PENNSYLVANIA STATE UNIVERSITY
UNIVERSITY PARK, PA. 16802

AN ASSESSMENT INVENTORY FOR THE DAY CARE CHILD

Volume II

FIELD EVALUATION AND PRELIMINARY FINDINGS

Center for Human Services Development

College of Human Development

The Pennsylvania State University

University Park, Pa. 16802

Center for Human Services Development

AN ASSESSMENT INVENTORY FOR THE DAY CARE CHILD[1]

Volume II

FIELD EVALUATION AND PRELIMINARY FINDINGS

Elizabeth P. Kirchner

The Pennsylvania State University

ABSTRACT

This is the second of a two volume report on an inventory of measures designed to assess day care children's understanding of their physical, social and personal world. It presents evaluations of the measures based upon findings from 282 day care children. Also presented, where possible, are interpretations of the findings in terms of their contributions to understanding various aspects of the development of children during the preschool period.

---

[1]This report was prepared under contract with the Pennsylvania Department of Public Welfare. The opinions and recommendations expressed are those of the author and do not necessarily reflect the official position of the sponsoring agency.

ACKNOWLEDGEMENTS

Many staff members of the Pennsylvania Day Care Study Project have contributed their talents and expertise to the work reported in this volume. It is a pleasure to extend my sincere appreciation for their help.

The support of Donald L. Peters, Ph.D., Project Head, has been continuously reinforcing. His encouragement has been greatly appreciated.

Robert F. Heckard and Paul D. Mowery, statisticians, provided their professional expertise in the selection of data analysis techniques and the management of all statistical procedures. I am most grateful for their competence and assistance.

Some findings of the work reported in this volume were prepared for presentation in professional journals and meetings. Such preparation was aided greatly by the collaboration of Sarah I. Vondracek, Ph.D. The discussions of vocational aspirations and perceived sources of esteem in this volume reflect her contributions.

The assistance of Ted G. Kastelic in the analysis of the Social Roles component is also gratefully acknowledged.

Eileen P. Hahn, senior secretary, Ruth R. Brooks, and D. Lynette Steck have given generously of the able secretarial and interpersonal skills that smooth the path of research and research reports. Their excellent work is sincerely appreciated.

Table of Contents

TABLES

FIGURES

AN ASSESSMENT INVENTORY FOR THE DAY CARE CHILD

Volume II

FIELD EVALUATION AND PRELIMINARY FINDINGS

Elizabeth P. Kirchner

The Pennsylvania State University

CHAPTER 1

INTRODUCTION

Volume II of the report An Assessment Inventory for the Day Care Child presents the findings of the field evaluation of the Day Care Inventory, which was developed by the Pennsylvania Day Care Study Project. Briefly, the Day Care Inventory was developed to provide a means for investigating day care children's understanding of their physical, social, and personal world. Although it is individually administered, the Inventory was not developed for purposes of assessing individual children. It was intended, instead, to provide a means for assessing the interaction of day care, family, and personal variables in the development of children in general. The objectives and goals of day care operators and experts were key considerations in selecting those aspects of development for which measures were provided (Peters, Hendrickson, Marcus, & Ridley, 1972).

Purpose and Format

The primary aim of this volume is to present an evaluation of the Day Care Inventory and to make recommendations for changes, deletions, and additions so that more satisfactory measures will be available with

which to study the impact of day care on several aspects of a child's development.

The results discussed in this volume may also provide an initial understanding of the status, at one point in time, of the day care children who comprised the sample. Findings will be utilized in this and in subsequent reports of the Pennsylvania Day Care Study Project to provide a more specialized study of specific aspects of the day care child's functioning as it relates to his day care setting and his family and personal characteristics.

The Day Care Inventory has 16 assessment components and a face sheet. Fourteen of the components are direct assessment measures. These components are:

Dog and Bone	Social Roles
Unusual Uses	Social Competence Questions
The Many-Splendored Cube	Who Likes You?
Picture Naming	Good Person and Bad Person
Picture Interpretation	Body Parts
Gross Motor Skills	Wish
Visual Discrimination	Camera Game

Two additional components are utilized as indirect assessment measures. They are: Child Ratings, completed by the day care teacher, and Examiner Observations, completed by the psychometrist who administers the Inventory to the child. The face sheet, completed by the teacher and/or caseworker, provides personal and demographic data for each child.

The task of evaluating the various components of the Inventory has several aspects. These aspects can be broadly divided into issues

relating to the reliability and validity of each measure.[2] It is obvious that the data of this study do not allow for a complete appraisal of either the reliability or the validity of any of the measures. This situation is to be expected, however, during the early development of any measure.

In order to evaluate the measures of this inventory more fully, further efforts are needed:

> More information must be gathered regarding to what extent these measures are "reliable," that is, to what extent they are consistent, accurate, and free of contaminating, extraneous factors that are unrelated to what is to be measured. While reliabilities for some components (those utilizing numerical sums of several items) can be estimated using present data, other components (those utilizing categories of response) will require further work.
>
> In order to determine whether the measures assess change, children must be assessed at two points in time rather than on only one occasion. Inasmuch as it is suspected that the greatest changes take place in the initial stages of a child's day care experience, it is important to plan the first assessment as close as possible to a child's entry into day care.
>
> To determine whether the measures reflect differences in child care environments, children who have been exposed to different sorts of preschool and day care experiences and children who have been in no program should be compared.

---

[2] The interested reader is referred to the following two references for discussions of the concepts of reliability and validity:

(a) American Psychological Association. Technical recommendations for psychological tests and diagnostic techniques. Washington: Author, 1954.

(b) Anastasi, A. Psychological testing. (3rd ed.) New York: MacMillan, 1968.

> Some "traditional" measures should be added to the Inventory for the sake of comparison. Of course, there are no well-known traditional measures in most of the domains assessed in the Inventory, or they would have been used in the first place. Where the measures used in the Inventory were chosen over traditional measures, it was because they were shorter, more appealing to the child, etc. The question of whether they yield data as good as that of traditional measures still remains, however, and is an additional reason for adding traditional measures to the Inventory.

Project activities continue in the above areas. It must be stressed, however, that this volume presents only the results of an initial evaluation of the Inventory components based on an administration of the Inventory (a) at a single point in time, (b) to six or fewer children from each of 51 day care centers, which have not been identified as to program emphases, (c) to no "control" groups, i.e., those with no group care experience, and (d) without the concomitant administration of more traditional measures.

For some measures evaluation can be more complete than for others. Sometimes this state of affairs is a function of the kind of measurement involved. Inventory component scores which are numerical in nature and represent ordinal and interval scales can be inspected by statistical procedures of evaluation which are not appropriate for the nominal data which other components yield. Further, for some components there is more data which can contribute to the understanding of the "goodness" of the measure than there is for others; for example, when there is more than one measure within the Inventory which assess similar behaviors, their interrelationships can be examined.

In this volume, as in Volume I, each component is separately discussed. The following format is used for each component:

Description of the Component - As an introduction to each component, and for the reader's convenience, a brief description of the component and how it was administered is presented.

Administration and Scoring - Here the psychometrists' feedback is reported. Their firsthand experience provides valuable, and in some cases the only, input into answering questions such as: "Do the children like this component?" "Do the children understand the directions?" "Are the directions or questions too wordy or too brief?"

When appropriate, the methods of scoring are presented in this section. Also presented are the interscorer reliabilities of the scoring systems utilized.

Findings[3] - This section presents and discusses those data which most directly concern the "goodness" of the measure beyond the first two considerations of smooth administration and accurate scoring. Possible age, sex, and race differences in the measure, the various specific relationships with other measures, and item analyses are considered when appropriate.

Where possible, this section also attempts to interpret the findings in terms of their contribution to the understanding of various aspects of the development of children during the preschool period.

It should be noted here that in the chi-square analyses, frequently utilized in analyzing the data of this report, the appropriate Yate's correction was used when the degree of freedom (df) was limited to one.

---

[3] Presented in the Appendix, as an aid to the nontechnical reader, is a glossary of statistical terms and symbols used in this report.

Summary and/or Recommendations - This section contains a summary of the findings when recapitulating seems necessary and presents any recommendations for changes in the component prior to undertaking further work with it.

Additional Information and Considerations Regarding the Sample

The major characteristics of the sample have already been described in Volume I. Presented here in this volume are (a) notes regarding two major variables, race and locale, involved in the analyses of the data, (b) a note regarding the representativeness of the sample, and (c) a further description of the sample in terms of parental employment and the number of parents present in the home.

The Race and Locale Variables

In Volume I the variables of race (black - white) and locale (urban - nonurban) were thoroughly confounded. The majority (88%) of black children in the sample were from urban areas. If analyses were performed using the total sample, apparent race differences might well be a function of locale differences, or vice versa. To avoid this confusing state of affairs, two decisions affecting all race and locale analyses were made:

a. Race (white - black) differences were analyzed among the urban children only.

b. Locale (urban - nonurban) differences were analyzed among the white children only.

In the analyses of the Inventory components significant findings related to the locale dichotomy were sufficiently infrequent as to be

regarded as primarily chance occurrences. Locale analyses, hence, are not reported in the discussions of the separate components. It is, however, included in the added information on the general characteristics of the sample.

It should be noted here that the locale dichotomy is best described as urban-nonurban rather than as urban-rural as in Volume I, since the latter classification actually represents a heterogeneous grouping of large and small towns as well as some, but few, bona fide rural areas.

Representativeness of the Sample

The field psychometrists, the people who gathered the data for this report, provided information and impressions which are critical in evaluating both the Inventory and the sample upon which the analyses are based. Their feedback regarding each Inventory component is presented when that component is discussed. Their feedback regarding the sample of children with whom they worked is presented here, however, for it is useful in further understanding possible biases in the sample.

The field psychometrists felt that, in addition to the possible bias introduced by utilizing only children for whom parental permission was obtained, further bias in sample selection may have been introduced within centers despite their best efforts to select among the available children at random. They felt that some teachers, by their subtle or frank efforts, may have led them to work with the "best" children and, to a lesser extent, with the "worst" children. The tendency to try to ensure that the "best" children be worked with seemed partly related to a teacher's feeling that she was being evaluated indirectly through the child, i.e., "If the child performs well, it means I am a good teacher." The less frequently noted attempt to have the "worst" children

worked with seemed related to a desire to "get him out of the classroom" or to a hope that the Inventory (endowed with magical diagnostic potential) might help identify the child's problem and its solution. Regarding the latter point, there were a few cases in which teachers felt, for example, that brain damage could be diagnosed via the Inventory. As one psychometrist put it, "I was the only professional they had seen, and they hoped I had all the answers."

To return to the matter of bias, it is likely that further biasing of the sample occurred when only a small sample of five or six children was selected at each center. This matter is less critical for evaluating the Inventory, the focus of this volume, than it is for evaluating day care and its variations, the focus of further study for which the Inventory is intended. In the latter case it is crucial that the sample of children be as representative as possible of all day care children. It is thus recommended that, to reduce the possibility of selection bias within a classroom in any further work, all children within a classroom be selected. Only those children for whom parental permission is denied should be excluded.

Further Description of the Sample: Parental Employment and Parents Present in the Home

Global characteristics of the total sample regarding employment and parents present in the home were described in Volume I. Presented here are further findings which help describe the sample in greater detail.

It should be noted that these comparisons are based upon incomplete information; for many children the desired information could not be

obtained. For example, employment status information was available for only about two-thirds of the sample, more for mothers and less for fathers. Therefore, when the race and locale subgroups were divided into specific occupational classifications, the number of cases became relatively small. The risk involved in making generalizations to the entire sample, and of course to the Pennsylvania day care population must therefore be stressed.

Table 1 presents the employment status of mothers and fathers for race and locale subgroups. It can be seen from Table 1 that there were

TABLE 1

RACE AND LOCALE COMPARISONS OF PARENTAL EMPLOYMENT STATUS

Group	Mothers			Fathers		
	Employed	Unemployed	Percent employed	Employed	Unemployed	Percent employed
Race comparison						
White urban	24	30	44	33	7	82
Black urban	42	16	72	37	3	92
	$\chi^2 = 7.92$, $df = 1$, $p = .005$			$\chi^2 = 1.03$, $df = 1$, $p = .31$		
Locale comparison						
Urban white	24	30	44	33	7	82
Nonurban white	40	59	40	65	10	87
	$\chi^2 = .10$, $df = 1$, $p = .75$			$\chi^2 = .10$, $df = 1$, $p = .75$		

no significant locale differences in maternal or paternal employment status and no significant race difference in paternal employment status. There was, however, a significant race difference in maternal employment status. More black than white mothers were employed (72% vs. 44%).

Table 2 presents the occupations of employed fathers according to race. Occupations are categorized in the table by the code used by the U. S. Bureau of the Census. Chi-square analyses of the data in Table 2 indicated that the occupational distributions were different for blacks and whites. Whites were more frequently represented in professional, skilled blue collar, and academic student categories, and blacks were more frequently represented in clerical and service occupations. Blacks also more frequently appeared in the nonspecific "other" category, a category utilized when the information did not specify type of occupation (e.g., "works at Ford").

Table 3 presents the occupational distribution of employed mothers as distributed by race. Again, in a chi-square analysis, the distributions were related. Inspection of Table 3 indicates that the largest difference appeared in the clerical category (whites = 50%, blacks = 24%) and in the nonspecific "other" category (blacks = 31%, whites = 4%).

The findings presented in Tables 1, 2, and 3 must be regarded as tenuous because they are based on limited information. They do, however, suggest that the black children came, more frequently than the white children, from homes in which the mother worked outside the home; also the occupations of the black children's fathers seem to have been more frequently specified according to where the work was done, or similar nonspecific descriptions, rather than according to the nature of the work done.

TABLE 2

OCCUPATIONS OF EMPLOYED FATHERS AS DISTRIBUTED BY RACE

Occupation	White urban		Black urban	
	n	%	n	%
Professional, teacher, etc.	7	21	4	11
Manager, official	0	0	2	5
Clerical	0	0	4	11
Sales	2	6	1	3
Craftsman, foreman	7	21	1	3
Machine & equipment operative	3	9	4	11
Service-worker	0	0	5	14
Laborer	2	6	1	3
Student - college, high school	3	9	0	0
Student - trade school, on the job	2	6	3	8
Other	7	21	12	32
Total	33	99	37	101

$\chi^2 = 21.49$, $df = 1$, $p = <.02$

Note.--The percentage totals do not equal 100 due to rounding errors.

TABLE 3

OCCUPATIONS OF EMPLOYED MOTHERS AS DISTRIBUTED BY RACE

Occupation	White urban		Black urban	
	n	%	n	%
Professional, teacher, etc.	5	21	6	14
Clerical	12	50	10	24
Sales	1	4	0	0
Machine & equipment operative	0	0	2	5
Service worker	5	21	9	21
Student (trade school, on-the-job)	0	0	2	5
Other	1	4	13	31
Total	24	100	42	100

$\chi^2 = 12.74$, $\underline{df} = 6$, $\underline{p} = .05$

The distribution of occupations by locale subgroups is presented in Table 4 for fathers and Table 5 for mothers. Chi-square analyses of these distributions show a significant relationship between locale and occupational distribution for both mothers and fathers. Inspection of Table 4 indicates a greater percentage of nonurban whites in laborer and service worker classifications and a smaller percentage in the nonspecific "other" category. The employed urban mothers, as seen in Table 5, were more frequently in professional and clerical positions and less frequently in service worker occupations than their nonurban counterparts.

TABLE 4

OCCUPATIONS OF EMPLOYED FATHERS AS DISTRIBUTED BY LOCALE

Occupation	Urban white		Nonurban white	
	n	%	n	%
Professional, teacher, etc.	7	21	6	9
Farmer, farm manager	0	0	1	2
Manager, official	0	0	2	3
Clerical	0	0	3	5
Sales	2	6	0	0
Craftsman, foreman	7	21	15	23
Machine & equipment operative	3	9	10	15
Service worker	0	0	5	8
Farm laborer & foreman	0	0	1	2
Laborer	2	6	14	22
Student: High school & college	3	9	0	0
Student: Trade school, on-the-job	2	6	0	0
Other	7	21	8	12
Total	33	99	65	101

$\chi^2 = 27.28$, $df = 12$, $p = <.01$

Note.--Percentage totals do not equal 100 because of rounding errors.

TABLE 5

OCCUPATIONS OF EMPLOYED MOTHERS AS DISTRIBUTED BY LOCALE

Occupation	Urban white		Nonurban white	
	n	%	n	%
Professional, teacher, etc.	5	21	1	2
Manager, official	0	0	1	2
Clerical	12	50	8	20
Sales	1	4	0	0
Craftsman, foreman	0	0	3	8
Machine & equipment operative	0	0	3	8
Private household worker	0	0	4	10
Service worker	5	21	20	50
Student (trade school, on-the-job)	0	0	1	2
Other	1	4	1	2
Total	24	100	42	102

$\chi^2 = 20.76$, $df = 9$, $p = <.02$

Note.--The nonurban percentage total does not equal 100 due to rounding errors.

It was noted earlier that the percentage of black mothers working outside the home was significantly greater than that of white mothers. This finding could be related to a racial difference in the percentage of mothers who "have" to work because there is no father in the home. Table 6 indicates that such was not the case however. It presents the breakdown of two-parent and mother-only families by racial subgroups

and shows that there was no significant relationship between racial subgroup and parents in the home.

TABLE 6

RACE COMPARISONS OF TWO-PARENT AND MOTHER-ONLY FAMILIES

Parents present in home	White urbans		Black urbans	
	n	%	n	%
Both parents	35	54	49	61
Mother only	30	46	31	39
Total	65	100	80	100

$\chi^2 = .53$, $df = 1$, $p = .47$

Next, for cases in which appropriate data were available, it appears that maternal employment was not related to parents-in-the-home for either white or black children. Table 7 presents the maternal employment status findings. It can be seen that the percentage of employed mothers was not significantly related to whether or not the father was present in the home.

TABLE 7

MATERNAL EMPLOYMENT STATUS AS RELATED TO PARENTS-IN-THE-HOME

Parents in the home	Mother employed		Mother unemployed	
	n	%	n	%
White urbans				
Both parents	12	44	15	56
Mother only	12	48	13	52
$\chi^2 = .000$, $df = 1$, $p = .98$				
Black urbans				
Both parents	24	69	11	31
Mother only	18	78	5	22
$\chi^2 = .257$, $df = 1$, $p = .61$				

CHAPTER 2

RATING SCALES

Child Ratings

Ratings of each child were made by his teacher. The 43 items of this rating scale are presented here for the convenience of the reader. The first 40 use the same type of rating. They are as follows:

1. Tries to be with another child or with a group.
2. Keeps trying even if something is hard to do.
3. Prefers to be by himself; wants to be let alone.
4. Gets in a temper if he can't have his way.
5. Loses interest and doesn't finish a puzzle, game or painting.
6. Is kind and sympathetic to someone who is upset or in trouble.
7. Likes to take part in activities with others.
8. Works a long time with a form board, puzzle, or other "achievement" toy, trying to complete it or get it right.
9. Watches others, but doesn't join in with them.
10. Gets impatient and unpleasant if he can't have his way.
11. Forgets a job or errand he started, as his mind wanders to other things.
12. Tries to make life easier for others; doesn't want to hurt them.
13. Enjoys being with others.
14. Pays attention to what he's doing; nothing seems to distract him.
15. Plays by himself rather than with others.
16. Pushes, hits, kicks others.
17. Gets distracted from what he's doing by what others are doing.
18. Is willing to share candy, food or belongings with others.
19. Seeks others out to get them to play with him or join in an activity.
20. Sticks to something he starts until it is finished.
21. Goes off by himself when others are gathering to dance or play together.
22. Gets angry when he has to wait his turn or share with others.
23. His attention wanders from what you're telling him.

24. Tries to help when he's asked.
25. Goes up to others and makes friends; doesn't wait for them to come to him.
26. Quietly sticks to what he's doing even when others are making noise or doing things nearby.
27. Tends to withdraw and isolate himself, even when he's supposed to be with a group.
28. Sulks, gets resentful, and won't do things he should.
29. Goes from one thing to another; quickly loses interest in things.
30. Awaits his turn willingly.
31. Follows directions and commands given by teacher.
32. Expresses his needs by words or gestures.
33. Has a hard time working with a group on a group activity.
34. Seems to notice feelings of others.
35. Uses imagination in his play.
36. Follows rules of behavior set by teacher.
37. Speaks soothingly, pats, or otherwise comforts a child who is hurt or unhappy.
38. Is curious to find out about things and people and events.
39. Tries to overcome obstacles and difficulties by himself.
40. Likes to have help even with easy things.

These items are rated according to the following scale: (1) almost never; (2) sometimes; (3) frequently; (4) almost always.

Three items, related to routine self-care, utilize other scale point definitions. They are:

41. Which best describes child's toileting behavior? (Please check)

_____ 1. Does not make needs known
_____ 2. Expresses needs but has occasional accidents
_____ 3. Expresses needs; avoids accidents, but needs some help
_____ 4. Uses toilet without help

42. How does child manage with dressing?

_____ 1. Cannot manage any clothing
_____ 2. Can manage most clothing except zipping and buttoning
_____ 3. Can dress self except for tying shoes
_____ 4. Can dress self completely

43. Can child wash face and hands?

_____ 1. Cannot wash face and hands
_____ 2. Can wash face and hands with help
_____ 3. Can wash face and hands without help

An additional item, which will be separately considered elsewhere, asks the teacher, "What three or four single words best describe this child to you?," and provides spaces for the teachers' descriptive words.

Administration

The psychometrists reported that the teachers were, for the most part, very cooperative in completing the Child Ratings form. They stressed, however, that the ratings may not always accurately reflect the child's behavior. Many teachers evidently interpreted an evaluation of the child as an evaluation of the teacher. Desiring to be judged a good teacher, a teacher may have, unwittingly or even knowingly, slanted her ratings in the direction of a "look good" or a socially desirable response.

It must be acknowledged, then, that this possible impediment to the validity of the ratings exists in the present sample to an unknown extent. This factor would operate to skew ratings spuriously in the good, socially valued direction. Another previously discussed factor, that of possible bias in the sample toward the "good" children being disproportionately represented, would have the same effect. Another possible source of bias (also previously discussed), the possible overinclusion of "worst" children, would influence the ratings in the opposite direction.

Findings

Distribution of Scores. The first 30 ratings of the scale are those of Schaefer and Aaronson (no date). Represented are six characteristics or traits, each comprised of five ratings. The six characteristics are extraversion, task orientation, introversion, hostility, distractibility, and considerateness. The distribution of the scores for these six characteristics are presented in Figure 1. Inasmuch as a score for a trait represents the sum of five ratings of from 1 to 4 each, the total score has a possible range of 5 to 20.

Thirteen supplementary ratings are single items specifically designed for the Day Care Inventory. The distributions of 10 of these items, which are rated from 1 to 4, are presented in Figure 2. Figure 3 presents the distributions of the remaining three supplementary items, which involve routine self-care skills.

It can be seen from Figures 1, 2, and 3 that the distribution of many ratings was considerably skewed, while some traits yielded more normally distributed ratings. Many of these traits can be assumed on theoretical bases to be normally distributed (a bell shaped distribution, mounding in the center and diminishing toward each extreme). The following are various reasons which might help to explain why some of the distributions departed from expected normality:

a. The rating judgments were strictly objective and reflect the behavior of the children, but the children in the sample represent a biased or skewed sample, hence the skewed ratings.

b. The rating judgments were not strictly objective and have been biased toward one extreme or the other by factors unrelated to the behavior of the children.

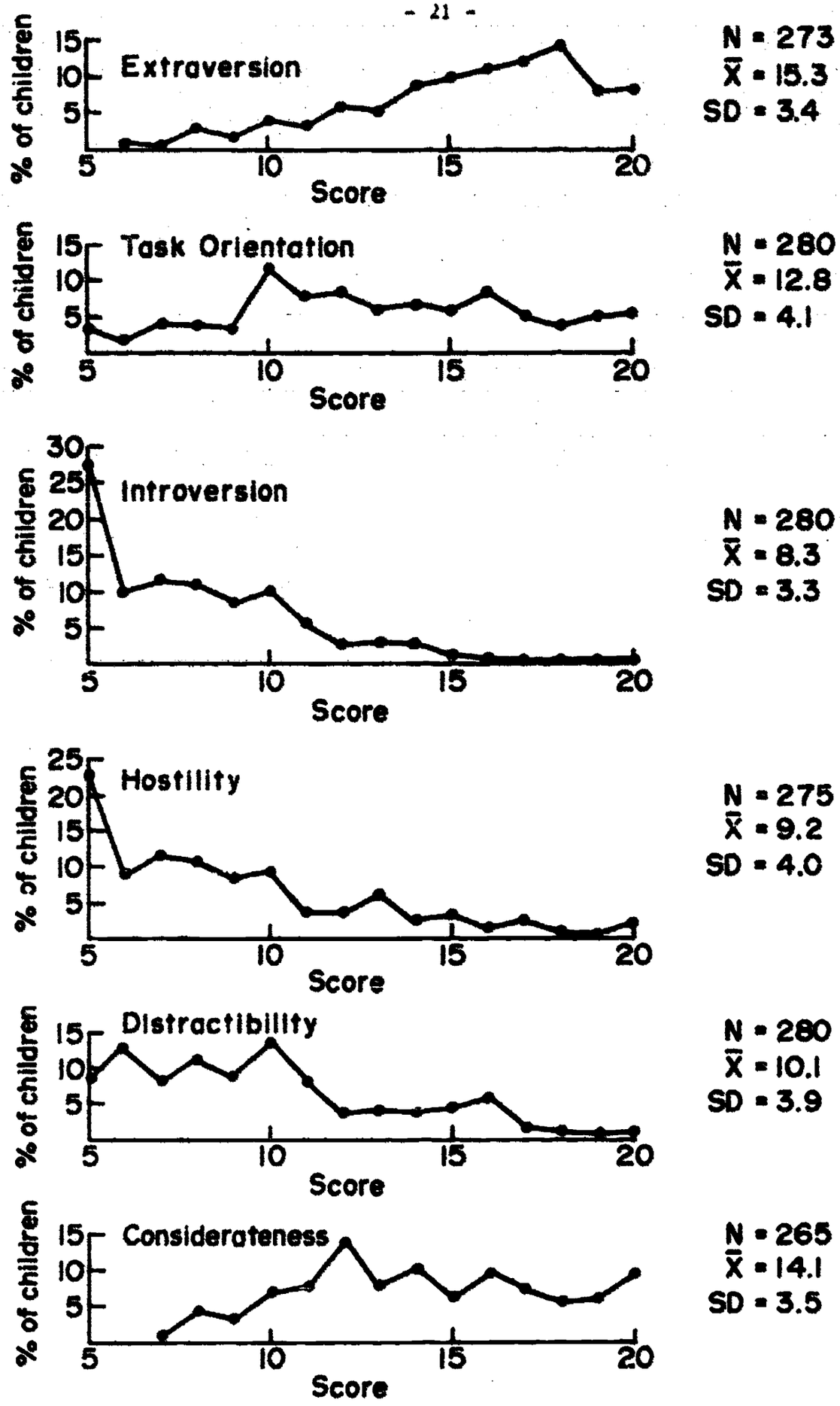


Figure 1. Distributions of the scores for the characteristics assessed by the Schaefer and Aaronson ratings.

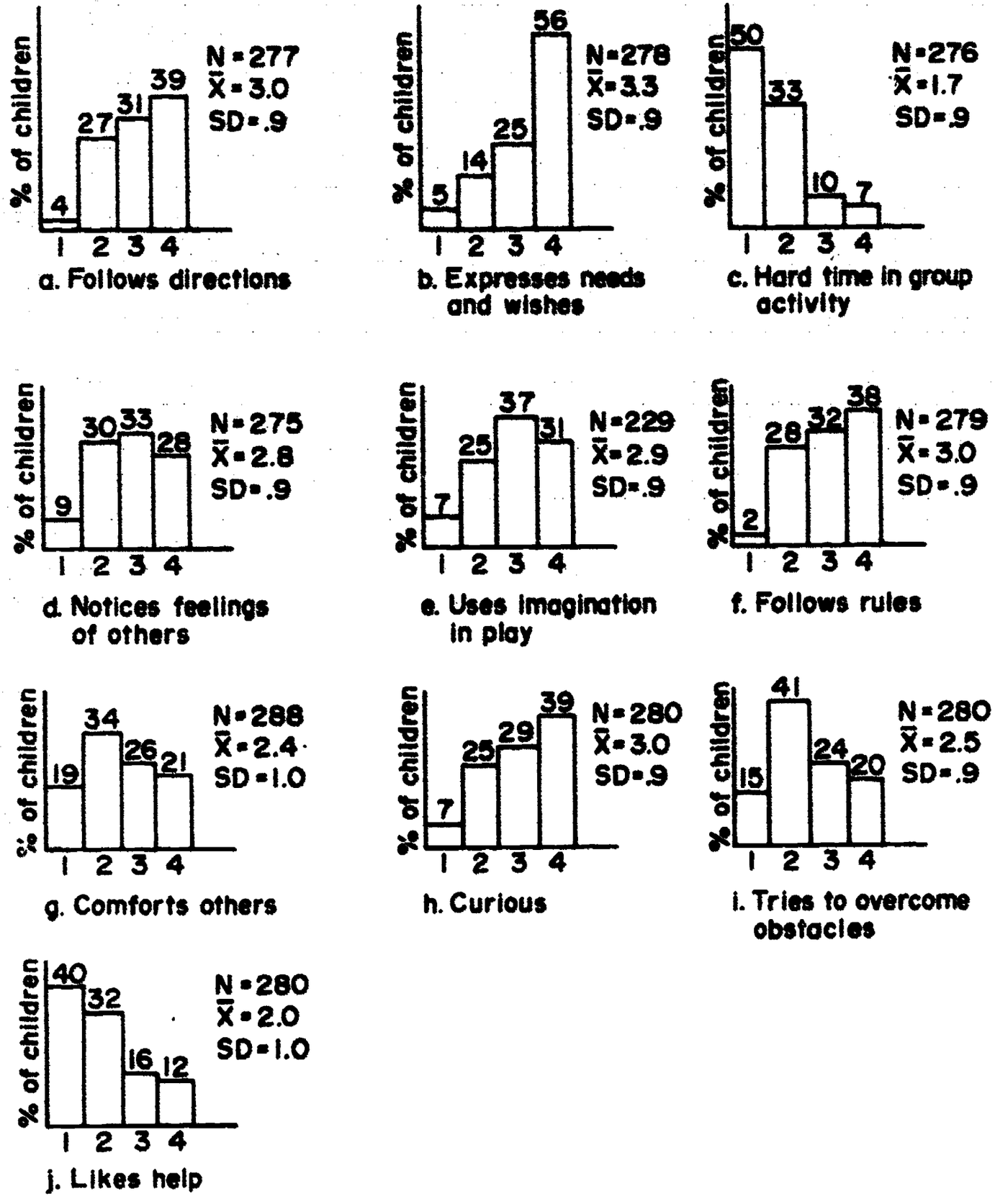
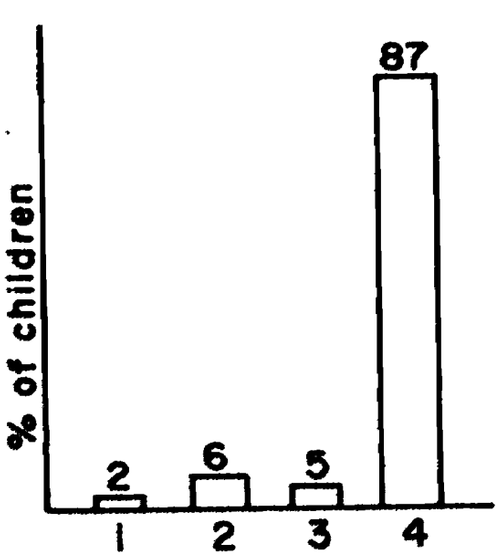


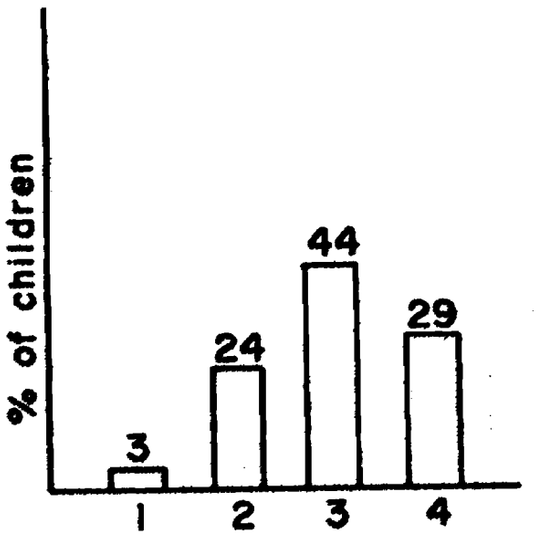
Figure 2. Distributions of the scores for ten of the supplementary ratings (1 = almost always, 2 = frequently, 3 = sometimes, 4 = almost never).



a. Toileting behavior

($\bar{X}$ = 3.8 SD = .7)

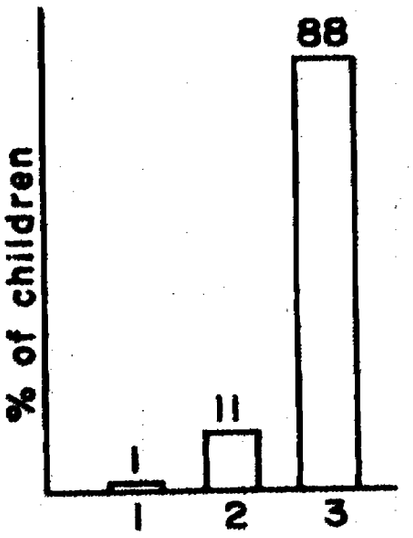
1. Does not make needs known
2. Occasional accidents
3. Needs some help
4. Uses toilet without help



b. Dressing behavior

($\bar{X}$ = 3.0 SD = .8)

1. Cannot manage
2. Manages except zipping & buttoning
3. Manages except tying shoes
4. Dresses self completely



c. Washing

($\bar{X}$ = 2.9 SD = .4)

1. Cannot wash face & hands
2. Can wash hands & face with help
3. Washes without help

Figure 3. Distributions of the scores for the three routine self-care ratings.

c. The items themselves do not adequately represent the traits or behaviors they are supposed to assess.

At present, it is not possible to identify the reasons for skewness which occurred in the ratings, although its presence suggests certain recommendations, presented at the end of the chapter, regarding the future use of this rating scale.

It is also recognized that markedly skewed distributions are hazards to straightforward data analyses. Many conventional statistical tests, such as the t-test for differences between means, assume normality of distribution. Conclusions based upon such tests are invalidated or vitiated to the extent that the distributions depart from normality. In some cases, however, alternate tests or mathematical data transformations are possible so that greater confidence can be placed in the analyses.

Subgroup Analyses of the Schaefer-Aaronson Scores. Means and standard deviations for the six Schaefer-Aaronson traits (extraversion, task orientation, intraversion, hostility, distractibility, and considerateness) are presented in Tables 8-13. Findings for the total group and for the age, sex, and race subgroups are given, together with the probability (p) of the subgroup differences being due to chance. From these tables it can be seen that the scores have satisfactory variability, very small standard deviations would have indicated that the scales were insensitive to differences among individual children.

The findings for the three age subgroups indicated, first of all, significant age differences in task orientation, an increase with age,

TABLE 8

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR THE EXTRAVERSION RATINGS

Group		N or n	$\bar{X}$	SD	p
All children		273	15.3	3.4	
Age:	3	65	14.4	3.7	
	4	135	15.6	3.3	.06
	5	73	15.3	3.2	
Sex:	Male	140	15.1	3.7	ns
	Female	133	15.5	5.4	
Race:	White	70	15.5	5.3	ns
	Black	83	16.1	3.9	

TABLE 9

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR THE TASK ORIENTATION RATINGS

Group		N or n	$\bar{X}$	SD	p
All children		272	12.8	4.1	
Age:	3	64	11.2	3.8	
	4	137	12.7	4.1	<.0001
	5	71	14.2	3.9	
Sex:	Male	138	12.3	7.7	ns
	Female	134	13.3	8.0	
Race:	White	70	13.6	7.3	ns
	Black	80	12.7	11.6	

TABLE 10

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR THE INTROVERSION RATINGS

Group		N or n	$\bar{X}$	SD	p
All children		270	8.3	3.3	
Age:	3	64	9.0	3.4	
	4	137	8.1	3.5	ns
	5	69	8.1	2.9	
Sex:	Male	137	8.5	7.6	ns
	Female	131	8.1	8.9	
Race:	White	69	7.6	9.3	ns
	Black	79	8.8	10.7	

TABLE 11

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR THE HOSTILITY RATINGS

Group		N or n	$\bar{X}$	SD	p
All children		275	9.2	4.0	
Age:	3	63	9.4	4.1	
	4	140	9.0	4.1	ns
	5	72	9.2	3.9	
Sex:	Male	140	9.8	5.8	<.10
	Female	135	8.5	7.3	
Race:	White	71	8.9	5.5	ns
	Black	83	9.5	8.4	

TABLE 12

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR THE DISTRACTIBILITY RATINGS

Group		N or n	$\bar{X}$	SD	p
All children		270	10.1	3.9	
Age:	3	64	11.6	4.2	
	4	135	9.8	3.7	<.01
	5	71	9.5	3.6	
Sex:	Male	139	10.7	7.3	.20
	Female	131	9.5	8.8	
Race:	White	69	9.6	8.8	ns
	Black	81	10.4	9.7	

TABLE 13

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR THE CONSIDERATENESS RATINGS

Group		N or n	$\bar{X}$	SD	p
All children		265	14.1	3.5	
Age:	3	64	13.7	3.2	
	4	134	14.3	3.7	ns
	5	67	14.3	3.6	
Sex:	Male	136	13.4	8.5	.18
	Female	129	14.9	9.0	
Race:	White	70	13.9	8.0	ns
	Black	77	14.6	11.5	

and in distractibility, a decrease with age. These differences are to be expected for the developing child, and hence they contribute to confidence in the validity of these scales. Two other traits, extraversion and considerateness, showed age trends which could also be expected, but the differences among the groups did not reach the .05 level of statistical significance. The remaining traits, introversion and hostility, showed no age-related differences. Inasmuch as hostility items reflect hitting behavior, overt display of temper, etc., one would have expected a decrease in the hostility ratings with age. The hostility ratings, however, were very low to begin with, suggesting that, among many possible alternatives, (a) the sample children may have been a particularly "nonhostile" group or (b) a display of hostility is relatively infrequent for all day care children and not related to the age of the child.

Sex differences in these ratings showed only one statistically significant difference: males were rated higher in hostility than females. This finding is consistent with considerable previous research. Trends toward sex differences in distractibility (males higher) and in considerateness (females higher) were also found and are consistent with expectations. Ratings for the remaining three traits (extraversion, intraversion, task orientation) indicated no tendency toward statistically significant sex differences.

No race differences were expected or found for these characteristics.

Subgroup Analyses of Scores on the Supplementary Ratings. The thirteen single supplementary ratings were each analyzed by the chi-square technique for possible age, sex, and race differences. A

summary of the significant findings and trends appearing in these analyses is presented in Table 14.

TABLE 14

SUMMARY OF COMPARISONS WITHIN THE SUBGROUPS FOR THE SUPPLEMENTARY RATINGS

Rating (abbreviated in some cases)	p	Direction of difference
Age comparisons		
Dresses self	<.001	3<4<5
Toilets self	<.01	3<4=5
Washes face and hands	<.01	3<4<5
Likes help	<.01	3=5<4
Overcomes obstacles by self	.08	3<4<5
Notices others feelings	.10	3<4<5
Comforts other child	.11	5<4<3
Sex comparisons		
Follows rules of behavior	<.001	males<females
Comforts other child	<.001	males<females
Notices others feelings	.04	males<females
Follows directions of teacher	.09	males<females
Race comparisons		
Is curious	.02	whites<blacks
Uses imagination in play	.06	blacks<whites
Comforts other child	.07	whites<blacks
Overcomes obstacles by self	.08	blacks<whites
Dresses self	.08	blacks<whites

Regarding age differences, four items showed significant results. Ratings on the three self-help items (dressing, toileting and washing) increased with the age of the child. On the fourth item, likes to have help even with easy things, the 4-year-olds were rated as more independent than the 3- or 5-year-olds, who were not rated as significantly different. Three trends toward significant age differences appeared. One suggests an increasing ability to overcome obstacles self-reliantly. The remaining two trends show a paradoxical comparison. On the one hand noticing the feelings of others seems to increase with age, but on the other comforting a hurt or unhappy child seems to decrease with age.

Sex differences appeared in relation to girls' greater compliance with authority. They had a significantly higher rating for following rules of behavior set by the teacher and showed a trend toward a higher rating for following directions and commands given by the teacher. Noticing others' feelings and comforting a hurt or unhappy child were also behaviors rated higher for girls than for boys, a trend which is in keeping with societal sex roles. These sex differences are consistent with previous research (see Mussen, 1970).

Race differences produced only one finding significant beyond the .05 level. Blacks were rated as more curious to find out about things, people, and events. A trend appeared, however, in their being less imaginative in their play. Two trends related to dependence on others for help (overcoming obstacles and dressing self) suggested that black children more frequently need or accept such help. A final trend indicated that black children more frequently comfort a hurt or unhappy child.

The findings from these supplementary single item ratings should be regarded as interesting leads for further investigation. Single items themselves are rarely of sufficient reliability to warrant firm confidence in results based upon them.

Examiner Observations

The psychometrists used the Examiner Observations rating scale to rate each of the children they worked with on 13 bipolar items which tap a child's response to assessment situations. Each item on the Examiner Observations scale has a 4-point rating system which omits the neutral midpoint and is defined by its polar opposites. The items are as follows:

1.	Came with examiner readily	vs.	Very resistant to coming with examiner
2.	Friendly and outgoing	vs.	Reserved, shy
3.	Comfortable throughout testing	vs.	Ill at ease throughout testing
4.	Cooperative, tries to follow directions	vs.	Negativistic refusal
5.	Pays close attention	vs.	Easily distracted
6.	Persistent	vs.	Gives up easily or can't give up (specify by underlining)
7.	Eager to continue	vs.	Seeks to terminate
8.	Challenged by hard tasks	vs.	Prefers only easy tasks
9.	Speed of response very rapid	vs.	Speed of response very slow
10.	Speech: clear, easy to understand	vs.	Extremely difficult to understand: mostly unintelligible
11.	Talkative: verbalizes a good deal	vs.	Taciturn: talks very little

12. Motor activity: overactive vs. Underactive

13. Guessing: a great deal vs. Resists guessing

Some of the items not only rate a specific behavior but also combine to assess a more global characteristic. The sum of the first four yield a social competence score, and the sum of the second four yield a task competence score.

Administration

For the most part the psychometrists felt able to utilize the Examiner Observations ratings without difficulty. They stressed that these ratings must be made immediately after working with a child while impressions are fresh. Critiques of two items were specified: the "challenged by hard tasks" item was identified as an inappropriate one in a case where no task is hard for the child, and the "speed of response" item did not specify the type of responses upon which to base judgment.

Findings

Distribution of Scores. The distributions of the Examiner Observations scores are presented in Figures 4, 5, and 6. Figure 4 shows the distributions for the first four ratings together with the distribution of the social competence score (which is obtained by summing these four ratings). Figure 5 presents the distributions of the task orientation ratings: the four component ratings and the global task orientation rating of which they are components. Distributions of the five additional ratings are presented in Figure 6.

The great majority of the distributions were markedly skewed in the "best" direction. Only three (speed of response, motor

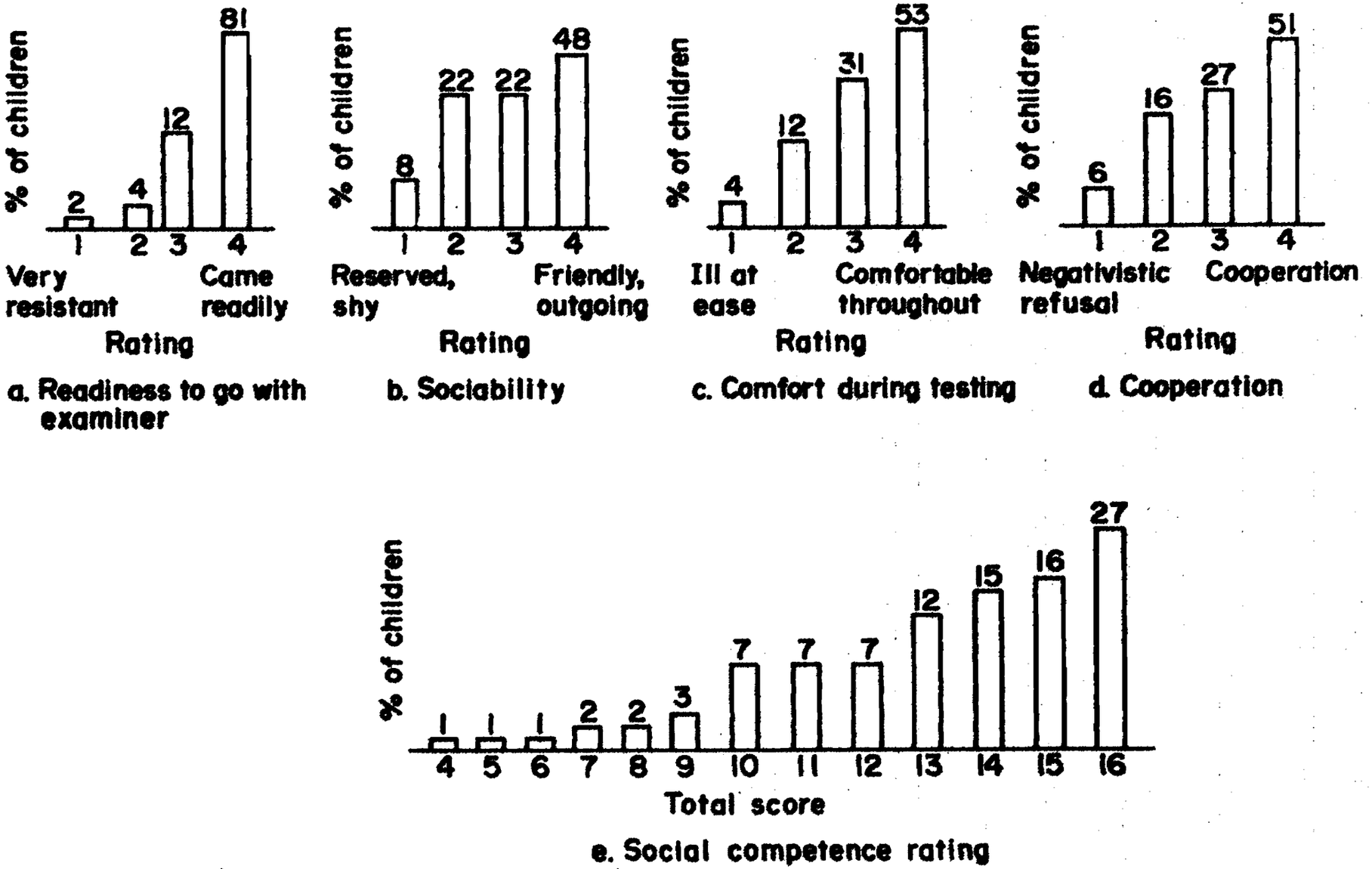


Figure 4. Distribution for the social competence rating and its four components.

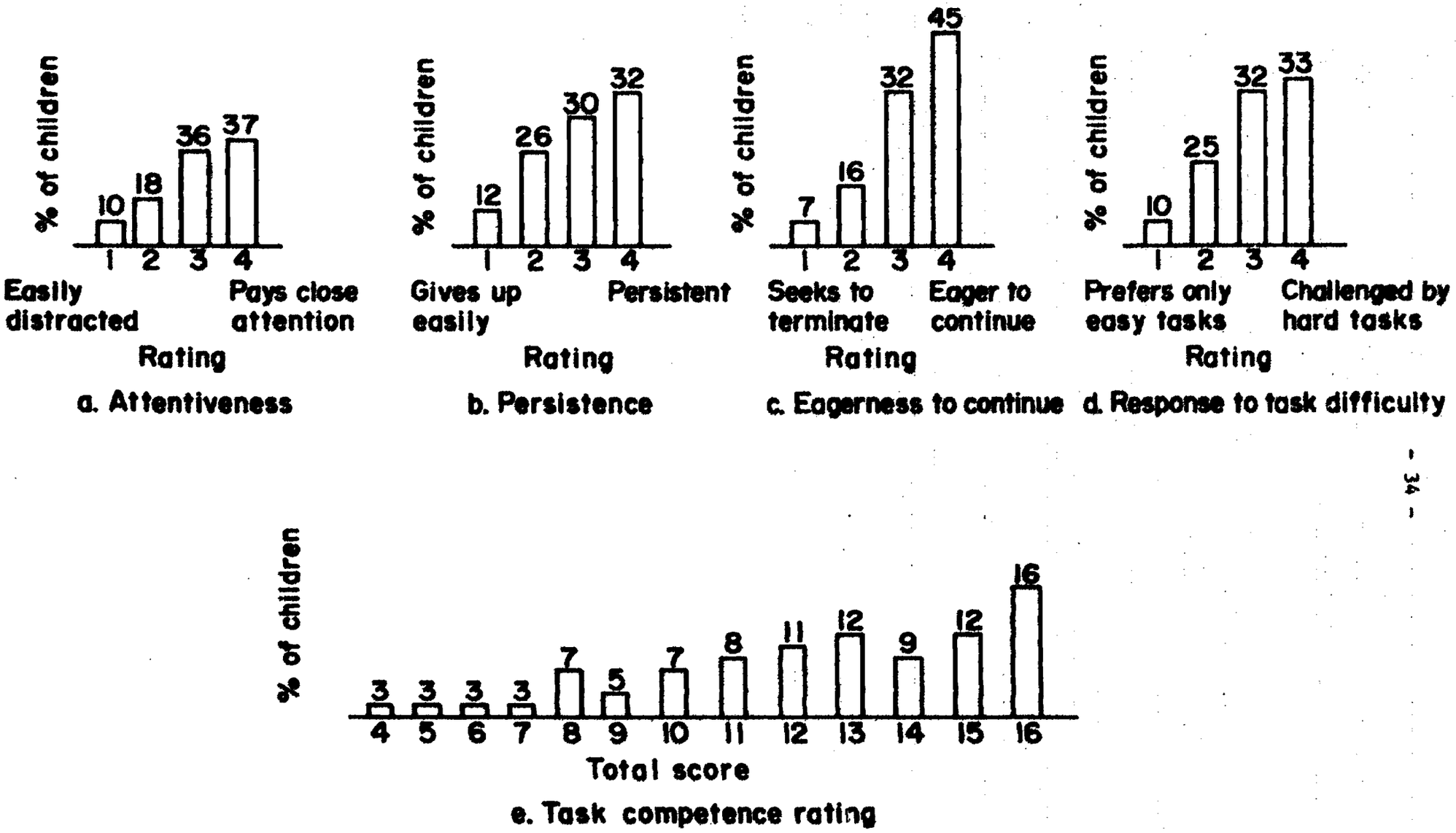


Figure 5. Distribution for the task competence rating and its four components.

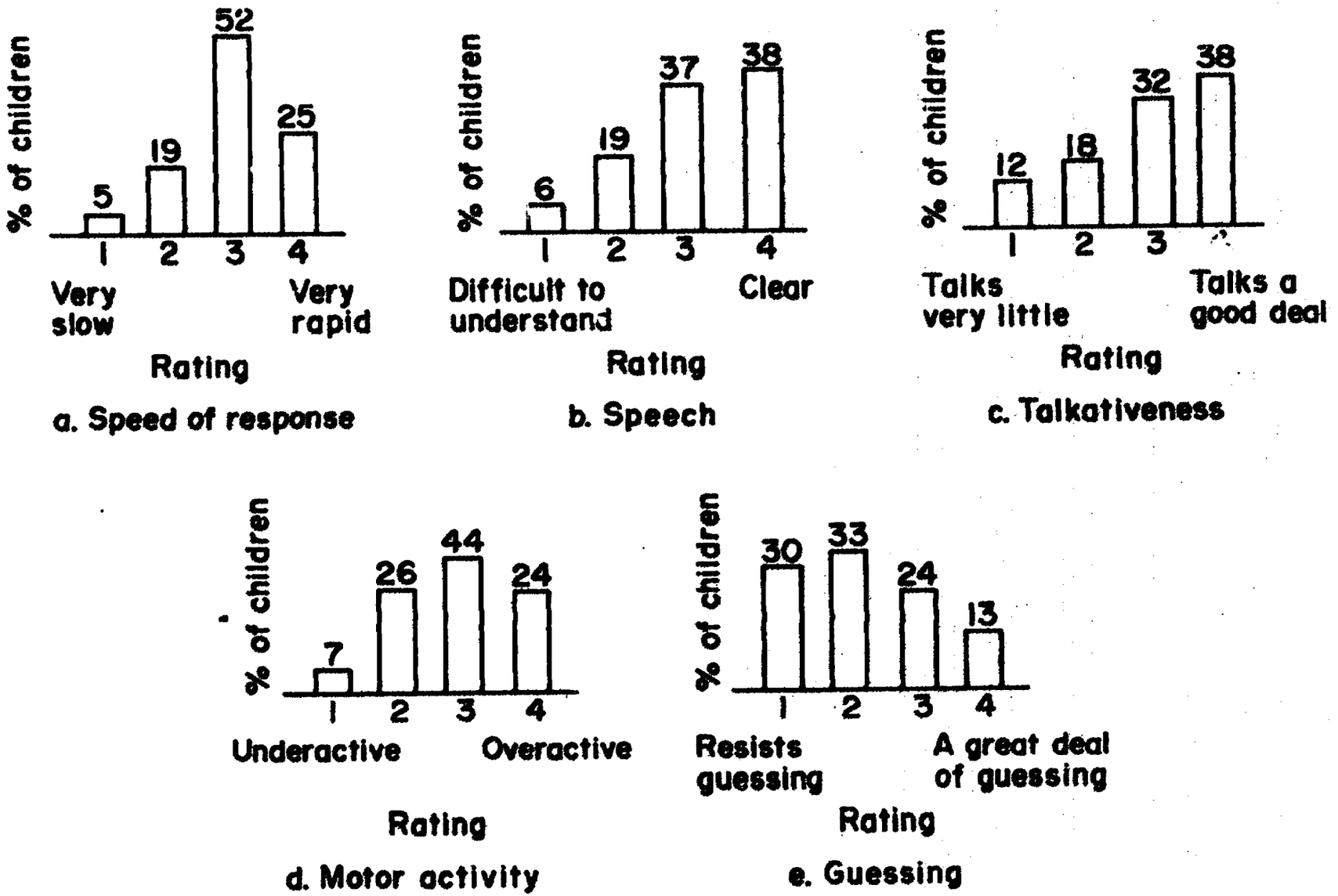


Figure 6. Distributions for the last five items on the Examiner Observations rating scale.

activity, and amount of guessing) did not indicate this skewness. The possibility grows more firm that the sample was saturated with children who were socially skilled and able to function well in a testing situation. Other possibilities exist, of course, but seem less likely. It might be suggested, for example, that the examiners' skill and ability led to these markedly "successful" ratings or that, in their desire to be judged good psychometrists, they knowingly or unknowingly inflated their ratings. It is my opinion, however, that the professional psychometrists were mature, experienced, and skilled to the extent that there is little reason to accept the latter possibility and that no amount of skill on the part of psychometrists could lead to these sorts of ratings for a normal distribution of children.

Age, Sex and Race Comparisons. Because of the skewness in the distributions subgroup comparisons for each examiner observation and the two global scores were based on chi-square analyses of the scores rather than on t-tests and F-tests for differences in the mean scores. Table 15 presents a summary of statistically significant subgroup comparisons (those with a probability of less than .05) and of trends (in this case, those findings with a probability between .05 and .20). These two sorts of findings are grouped separately for each subgroup comparison. Inspection of Table 15 indicates that a large number of ratings showed age differences, while sex and race subgroup comparisons yielded fewer significant differences. The findings regarding age differences impart some confidence in the ratings; one "expects" these findings for the most part. The three ratings which did not show age-related differences (comfortable throughout testing, talkative, and

motor activity) might also have been expected to show age differences but did not.

TABLE 15

SUMMARY OF COMPARISONS WITHIN THE SUBGROUPS

FOR EXAMINER OBSERVATIONS

Rating name (abbreviated in some cases)	p	Direction of difference
Age comparisons		
Cooperative	<.001	3<4<5
Pays close attention	<.01	3<4<5
Persistent	<.01	3<4<5
Eager to continue	<.01	3<4<5
Challenged by hard tasks	<.01	3<4<5
Speech easy to understand	<.001	3<4<5
Guessing	<.02	3<4=5
Social competence	<.001	3<4<5
Task competence	<.02	3<4<5
Came readily	.12	3<4<5
Friendly and outgoing	.12	3<4<5
Speed of response	.09	3<4=5
Sex comparisons		
Speed of response	.17	males<females
Eager to continue	.07	males<females
Race comparisons		
Pays close attention	.05	blacks<whites
Speech clear	<.001	blacks<whites
Cooperative	.19	blacks<whites
Talkative	.10	blacks<whites
Overactive	.14	whites<blacks

Sex differences were scant. Only two trends appeared: males were less eager to continue and less rapid in response. More sex differences

could have been expected on the basis of the girls' supposed edge in social skills, verbal facility, etc., at these ages.

Race comparisons, about which it was difficult to make any prior predictions, yielded some differences which suggest that the black children were somewhat harder to work with (poorer attention, speech harder to understand, etc.). Further analyses may determine whether racial mix and match in child-examiner dyads relate to these findings.

Recommendations

For further work with Child Ratings and Examiner Observations the following recommendations are made:

a. In further work, all children from any one classroom in the sample should be included to the extent allowed by parental permission. Possible biasing toward overinclusion of "cooperative" children in the sample would therefore be reduced.

b. Where possible, a Child Ratings form should be filled out by two teachers who know the child well. Validity and reliability of ratings may thereby be increased.

c. For the Child Ratings the definition of the scale points should be changed. The current frequency-related scale (1 = almost never, to 4 = almost always) should be replaced with one focusing on how descriptive or characteristic the item is for the child. This latter scale was suggested by Aaronson in a recent personal communication: 1 = not at all like, 2 = very little like, 3 = somewhat like, 4 = very much like. It reportedly has been used extensively by Aaronson and Schaefer, and, being more successful psychometrically, has replaced their earlier frequency-related scale, which was adopted for the Inventory.

d. In Examiner Observations two ratings, speed of verbal response and speed of motor response, should be substituted for the ambiguous rating of speed of response.

CHAPTER 3

DOG AND BONE

The Dog and Bone component of the Day Care Inventory is a measure designed by Banta (1970) to assess a child's ability to generate innovative, alternate solutions to problems. The child is seated at a board with a house near each corner. A toy dog is placed on the child's side of the board; directly across from the dog is placed a bone. The examiner points out the houses, dog, and bone and indicates with his finger two routes by which the dog can get to the bone. Upon completion of the demonstration, the examiner asks the child to "take the doggie and find another way, a different way, for him to get his bone." Ten such trials are presented to the child.

Administration

The psychometrists were enthusiastic in general about this measure and its placement as the first item to be administered. They felt it was a good icebreaker and that it was an enjoyable, nonthreatening task for the child. The nonverbal nature of the required response was cited as a distinct asset for an initial task in an inventory which requires verbal responses on the majority of the measures.

Regarding the instructions given the child, some psychometrists felt that the demonstration items should involve actually moving the dog to the bone, as the child is later asked to do. Current instructions, adapted directly from those of Banta, call for the examiner to demonstrate by moving his finger across the board.

One psychometrist questioned the need for 10 trials. He felt that a smaller number of trials might be sufficient for a child to demonstrate versatility in finding alternate paths while at the same time saving a child who is unable to achieve much response variability from repetition-produced frustration or boredom. Other psychometrists felt strongly that 10 trials were essential to bring out individual differences in the ability to do the Dog and Bone task.

Recording the child's response was reported to be easily done, but, "You had to know ahead what you were doing and you had to pay close attention to doing it right." Thus prior training and practice seems essential.

One psychometrist suggested that a smaller board might be provided, especially for younger children, some of whom he felt had difficulty reaching across the board. Other psychometrists reported they had been able to vary placement of the board (nearer the child or on the examiner's lap, for example) to overcome any difficulties related to the span of a young child's reach.

Finally, one examiner, who evidently had had a close call with some very active children, suggested a spare dog be provided in case of breakage.

Scoring

Banta's system for scoring the child's responses credits each different path produced by the child and gives additional credit for complex pathways involving crossovers and reversals in the route. This system was utilized in scoring the present data.

Another scoring system was also adopted, and the two are compared in the following section. The second system gives credit when either the dog's path or mode of movement is varied. This second scoring system also includes the modification of giving no extra credit for path crossovers and reversals, the judgment of which seemed difficult for scorers and is thus a possible source of some unreliability of scoring.

This modified scoring system was developed on the basis of observations during the pilot testing of Dog and Bone. It was noted at that time that children occasionally varied the movements of the dog in its progress to the bone (e.g., rolling, head over tail, jumping, bouncing on his head). The modified scoring system which credited these variants was based on the hypothesis that these behaviors also reflect innovative behavior.

The reliability of scoring was assessed by measuring the percent of agreement on the scoring of each trial in 25 protocols by two independent scorers. There was 92% agreement in the Banta scoring and 94% agreement in the modified scoring. This mode of assessing reliability (trial-by-trial) is a stringent one. The more usual mode, that of correlating total scores of the two scorers, yielded correlations near .98.

Findings

Distribution of Scores

Dog and Bone was administered to all 282 children. The total score ranged from 0 to 25 for both the Banta scores and the modified scores. For the Banta scores, the mean was 5.63 and the SD was 4.39;

for the modified scores the mean was 6.52 and the SD was 4.44. Thus, both scoring systems yielded good, and approximately equal, variability of scores. The distribution of both the Banta and modified scores are presented in Figure 7. These distributions are seen as somewhat skewed toward the low end of the scale; the bulk of scores were between 0 and 10, and few children achieved scores in the 11-25 range. Both scores thus had more "top" (range and differentiation among high scores) than "bottom" (range and differentiation among low scores).

Subgroup Analyses of Scores

Analyses of variance and t-tests were conducted to investigate possible age, sex, and race differences in total scores. Table 16 presents the findings related to these analyses of subgroup differences.

It can be seen from Table 16 that there were highly significant age differences in both scores ($F = 9.35$, $df = 2/279$, $p = <.005$ for the Banta scores; $F = 10.49$, $df = 2/279$, $p = <.005$ for the modified scores). Tukey's tests subsequent to these analyses indicated significant differences between each age group's mean (3<4<5 in each case). Inasmuch as innovative behavior can be expected to increase with age, these sturdy age-related findings are supportive of confidence in this measure.

There was no significant sex difference in the Banta scores ($t = 1.51$, $df = 280$, $p > .10$). The sex difference in the modified scores was significant ($t = 2.06$, $df = 280$, $p < .05$): the boys achieved a higher mean score than the girls. The issue of a possible sex difference in creativity-related abilities is a complex one, but the research literature suggests that when a sex difference is found, males are more

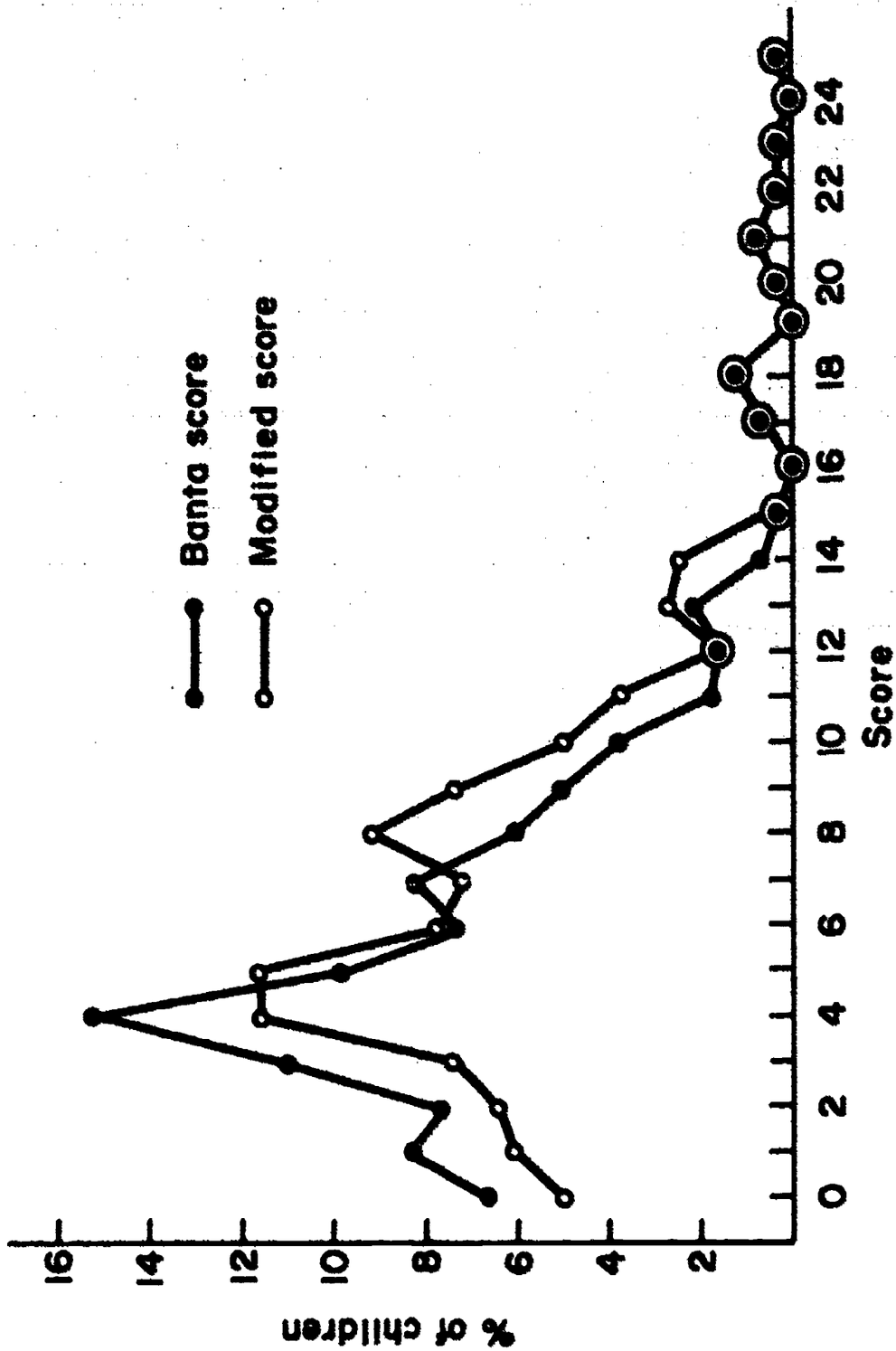


Figure 7. Distribution of the Banta and modified scores.

TABLE 16

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR

THE DOG AND BONE TOTAL SCORES

Group		N or n	Banta total score			Modified total score		
			$\bar{X}$	SD	p	$\bar{X}$	SD	p
All children		282	5.63	4.42		6.52	4.44	
Age:	3	66	3.89	3.41		4.70	3.76	
	4	142	5.70	4.44	<.005	6.58	4.39	<.005
	5	74	7.03	4.67		8.05	4.68	
Sex:	Male	143	6.04	4.76	>.10	7.06	4.76	<.05
	Female	139	5.21	4.00		5.97	4.11	
Race:	White urban	73	6.38	5.11	>.10	7.48	4.97	>.10
	Black urban	86	5.63	4.07		6.52	4.23	

frequently found to surpass females in these abilities. Thus, if any sex difference were to be predicted on the Dog and Bone component, the prediction would be in the direction of higher scores for males. From the findings obtained in this sample, the Banta scores would not support the prediction, whereas the modified scores would. The modified scoring system is thus seen as more sensitive to sex difference. Whether the sex difference reflected is in the realm of creative abilities is a

judgment which must await validation of the modified scores as representing creative behavior rather than other sex-related behaviors.

A final subgroup comparison concerns possible racial differences in the Dog and Bone scores. None were predicted and none were found.

Correlation Between the Banta and Modified Scoring Systems

Inasmuch as two scoring systems were utilized, it is appropriate to assess how closely correlated the two systems are. In addition to the findings of similar ranges of scores and the 1-point difference in mean scores of the two scoring systems shown in a previous section, the two systems yielded a correlation of .94. This correlation is very high and suggests that both scoring systems measure the same thing to a high degree. Ordinarily one type of score would be expected to "do as well" as the other, so it is somewhat surprising to find, as was presented in the preceeding section, that sex difference was found significant with the modified score and nonsignificant with the Banta score. It is tentatively suggested that whatever variance the two scores do not have in common is variance related to a sex difference of some sort.

Relationship of Scores to Selected Ratings of the Child

To shed light on validational issues, the Dog and Bone scores were analyzed in terms of their relationship to selected ratings the teachers made of the child's behavior in the day care center and to the ratings the examiner made of the child's behavior during the administration of the Inventory.

Of the six single-item ratings by the teacher which were selected for analyses, three were related to the issue of convergent validity

("Does the component measure what it says it measures?"). They involved behavior assumed to be similar, although by no means identical with, innovative behavior. These three ratings were (a) uses imagination in his play, (b) tries to overcome obstacles and difficulties by himself, and (c) is curious to find out about things and people and events. The three other ratings were related to divergent validity (indicating that a test is not related to what it should not be related to). They involved behaviors assumed to have no relationship to innovative behavior. These ratings were (a) follows directions and commands given by the teacher, (b) follows rules of behavior set by the teacher, and (c) likes to have help, even with the easy things. (It could also be argued that the last rating should be negatively related to innovative behavior.)

Because the child ratings were on a scale which ranged only 4 points and because many distributions were skewed to an extent to make correlational analyses undesirable, analyses of variance were utilized to explore the relationships between Dog and Bone scores and the six items selected from Child Ratings. The results are summarized in Table 17. It can be seen in Table 17 that the three ratings predicted to be related to Dog and Bone scores were found to be so and that the three ratings presumed to reflect behaviors irrelevant to the scores were indeed found to be unrelated to them. Inspection of the means in the cases where significant results were obtained indicates that children who were rated lower (1, almost never, and 2, sometimes) on uses imagination, is curious, and overcomes obstacles had lower Dog and Bone scores than did children who were rated higher (3, frequently, and 4, almost always). These findings contribute, albeit in modest fashion, to confidence in the validity of Dog and Bone as a measure

TABLE 17

SUMMARY OF ANALYSES OF VARIANCE FOR DOG AND BONE SCORES AND SELECTED CHILD RATINGS

Rating		df	Banta score			Modified score		
			$\bar{X}$	F	p	$\bar{X}$	F	p
Uses imagination:	1		4.6			4.8		
	2	3/225	4.3	3.55	<.01	5.3	4.93	<.01
	3		5.6			6.3		
	4		6.8			8.0		
Tries to overcome obstacles:	1		5.9			6.7		
	2	3/276	4.7	3.06	<.05	5.6	3.08	<.05
	3		6.4			7.4		
	4		6.4			7.3		
Is curious:	1		3.8			4.8		
	2	3/277	5.9	2.61	<.10	6.6	2.99	<.05
	3		5.0			5.8		
	4		6.3			7.3		
Follows directions:	1		5.2			6.8		
	2	3/273	5.4	.39	ns	6.4	.08	ns
	3		5.4			6.4		
	4		6.0			6.7		
Follows rules:	1		5.2			6.8		
	2	3/275	5.6	.03	ns	6.6	.02	ns
	3		5.6			6.5		
	4		5.7			6.5		
Likes to have help:	1		5.9			6.7		
	2	3/276	6.1	2.53	ns	7.0	2.29	ns
	3		4.0			5.0		
	4		5.6			6.7		

Note.--Ratings based on following scale: 1 = Almost never, 2 = Sometimes, 3 = Frequently, 4 = Almost always.

of innovative behavior. It is also noted that when the results via two scoring systems are compared, the modified system appears slightly more robust, yielding higher F's where significant relationships were predicted and lower F's where they were not.

Additional correlational analyses were performed to determine possible relationships between Dog and Bone scores and examiner ratings and other ratings made by the teacher. These are the examiner's ratings of task competence and social competence and the teacher's ratings of task orientation and distractibility. These ratings all represent sums of a number of component scales and are more fully discussed in Chapter 2. Banta reported correlations of Dog and Bone with examiner ratings of task competence and social competence based on scales similar, although not identical, to ours. He reported a correlation of .26 with social competence ratings and interpreted his finding as giving support to the view that the social psychological context of the testing situation is related to the child's producing creative solutions to problems. He found a correlation of .10 with task competence. Table 18 presents the findings from the present study's data. The major interpretation placed upon these results is that all relationships represented are either nonexistent or slight. Even the correlations which are significant in a statistical sense (a function of N) demonstrate minimal relationship in a real sense. There is thus little variability in Dog and Bone scores which is related to the behaviors reflected in these ratings by teachers and examiners. The strongest correlations are with social competence (.18 and .17), but they are lower than that reported by Banta (.26). While this particular finding may vitiate the interpretation referred to above, it is

TABLE 18

CORRELATIONS BETWEEN DOG AND BONE SCORES AND SELECTED RATINGS BY EXAMINER AND TEACHER

Rating	N	Banta score	Modified score
Ratings by examiner:			
Social competence	259	.18*	.17*
Task competence	257	.16*	.13
Ratings by teacher:			
Task orientation	261	.17*	.16*
Distractibility	270	-.10	-.09

*p <.05

felt that the totality of findings contribute additional support for the discriminant validity of the component.

Relationship with the Unusual Uses Component

Another component of the Inventory, Unusual Uses, was also selected because of its purported relevance to creative behaviors. Briefly, Unusual Uses, which is discussed in Chapter 4, is a measure which requires verbal response. Unusual Uses asks the child to give as many uses as he can for two common objects, a cup and a newspaper.

Since both the Dog and Bone and the Unusual Uses measure claim to "get at" creativity, their relationship was analyzed. The Unusual Uses total score correlates .21 with the Banta scores and .24 with the modified scores. These correlations indicate that the two measures, Unusual Uses and Dog and Bone, have little in common (about 5% of shared variance). However, they may each be assessing a quite different sort of creative behavior (generating innovative paths to a specific

spatial goal vs. generating innovative ideas regarding nonspecific purposes for items). Additionally, the mode of innovative response required (motoric vs. verbal) differs markedly.

Recommendations

This component should be retained in further studies, particularly those which wish to explore factors influencing innovative behavior. It is easily administered, is enjoyable for the children, can be reliably scored, and shows robust age-related differences. The modified scoring system seems to be somewhat more sensitive a measure than the original Banta system. Until there are more data upon which to base a choice, however, both scoring systems should be retained.

CHAPTER 4

UNUSUAL USES

The Unusual Uses component was included in the Inventory, as was Dog and Bone, as a measure of creativity. It calls for the generation of ideas regarding possible uses for two common objects, a cup and a newspaper.

Examiner instructions for "cup" are:

> You know what a cup is; you can use a cup for lots of things. Tell me all the things you can think of that you can use a cup for.

Instructions for "newspaper" are:

> Now think about a newspaper--you can use a newspaper for lots of things, too! Tell me all the things you can think of that you can use a newspaper for.

Administration and Scoring

The examiners reported neither particular difficulties with this component nor particular enthusiasm on their part or the children's. It seemed a straightforward, brief, "uninvolving" component.

The scoring procedure provided for tallying response frequency (e.g., drink milk, drink water, and dig in the sand with it), category frequency (e.g., drinking utensil, digging tool), and "original" responses. Despite other researchers' seeming ability to score originality, the attempts of this study were unsuccessful. A statistical criterion (that of a response appearing only once in 100 records) yielded too few responses for any analyses to be performed with them and yielded responses which two clinicians could not agree

upon as to their validity (measuring "originality" rather than "bizarrness which escapes originality"). The remaining scores yielded satisfactory reliabilities, as assessed by percent of agreement on two independent scorings of 75 protocols: cup frequency, 94%; cup category, 96%; newspaper frequency, 91%; and newspaper category, 95%.

Findings

Distribution of Scores

The Unusual Uses component was administered to 266 children. There are six scores to be considered: response frequencies for cup and newspaper and the total (combined) response frequency; and category frequencies for cup and newspaper and the total (combined) category frequency.

Figure 8 presents the distributions of response frequencies for cup, newspaper, and their total. It can be seen in Figure 8 that all distributions were skewed toward lower scores and that cup had a wider range of scores (0-11) than did newspaper (0-8). Further, almost one-half of the children (47.4%) gave only one response for newspaper, and one-third (33.1%) gave only one response for cup. The distribution of total cup and newspaper scores showed a continuous range of 0-12 (with one individual achieving a score of 17) and a less peaked distribution, which is a function of summing the two scores which make up the total. The mean response frequencies were: cup, 2.7 (SD = 2.0); newspaper, 1.6 (SD = 1.2); total 4.3 (SD = 2.9). The correlation between response frequencies on cup and newspaper was .55. This correlation indicates that there was a moderate relationship between the number of responses given to the two items.

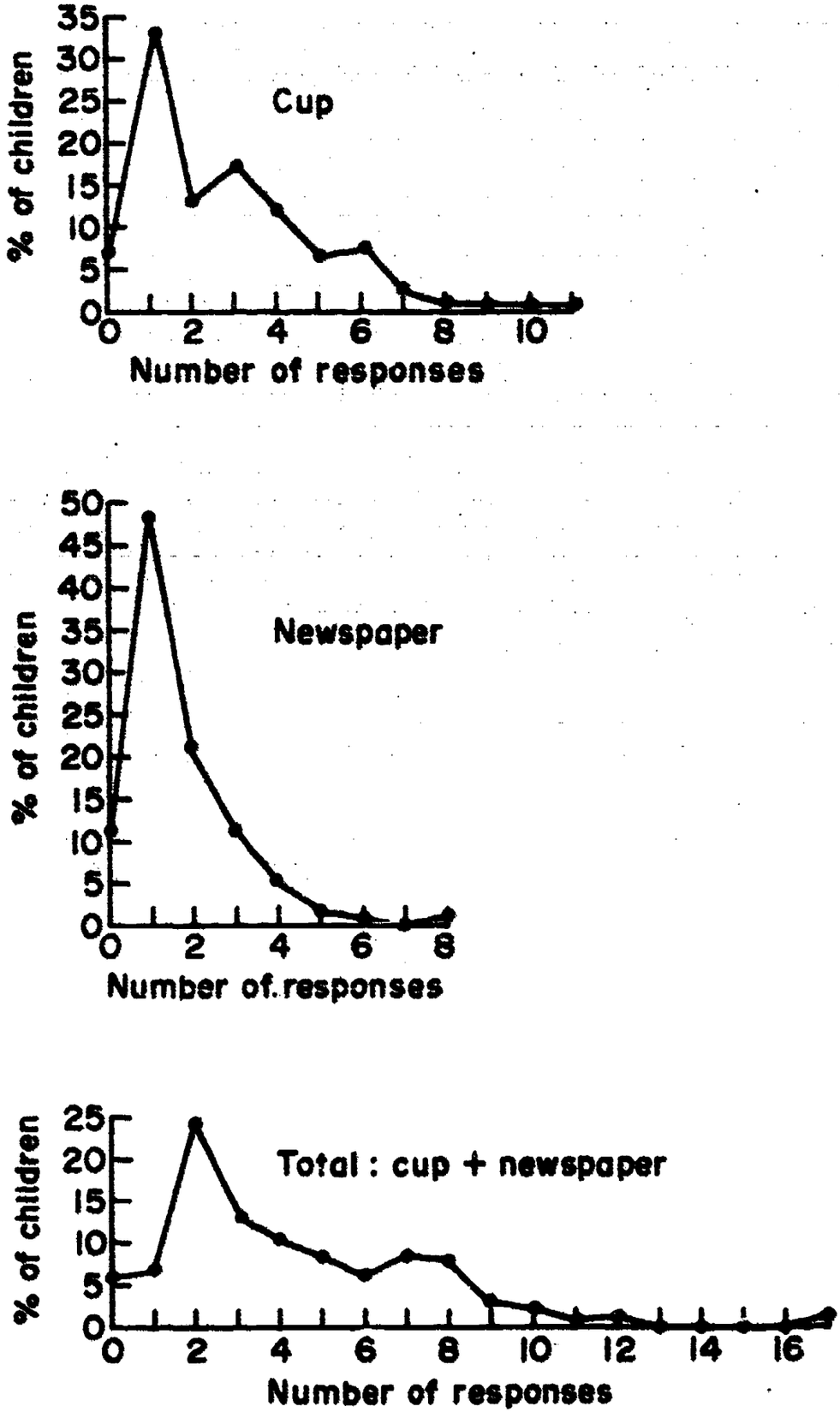


Figure 8. Distribution of response frequency for cup, newspaper and their combined total.

These findings regarding response frequencies indicate that, although there was a considerable range in the scores, on the average, there was only modest response from the children.

The distributions of category frequencies are presented in Figure 9. It can be seen that only one category was utilized by the majority of the children for cup (73.7%) and for newspaper (56.0%). A total category score of 2 was thus most frequent (47.4%). Again, there seems to be only modest responding represented in these scores.

Subgroup Analyses of Scores

Table 19 presents, for cup and newspaper frequency and the total of both, means and standard deviations for the total sample and the age, sex, and race subgroups. Table 20 presents findings for category frequency. Significance levels (p) for tests of possible subgroup differences are also presented in both tables. Of the subgroup comparisons made, age showed the only significant relationship to the scores; five of these analyses were significant, (response frequency: cup, $F = 14.72$, newspaper, $F = 11.52$, total $F = 17.34$, $df = 2/263$; category frequency: newspaper, $F = 5.46$, total $F = 5.66$, $df = 2/263$). Tukey tests showed a significant difference between each age (3>4>5) for these scores. A sixth score, category frequency for cup, was not significantly related to age.

No sex or race differences were predicted or found.

In general, these findings, inasmuch as they show the overall age progression which would be expected, support confidence in the measure. However, when the results of the present study are compared with those reported by Peters and Stein (1966) and by Gish (1971)

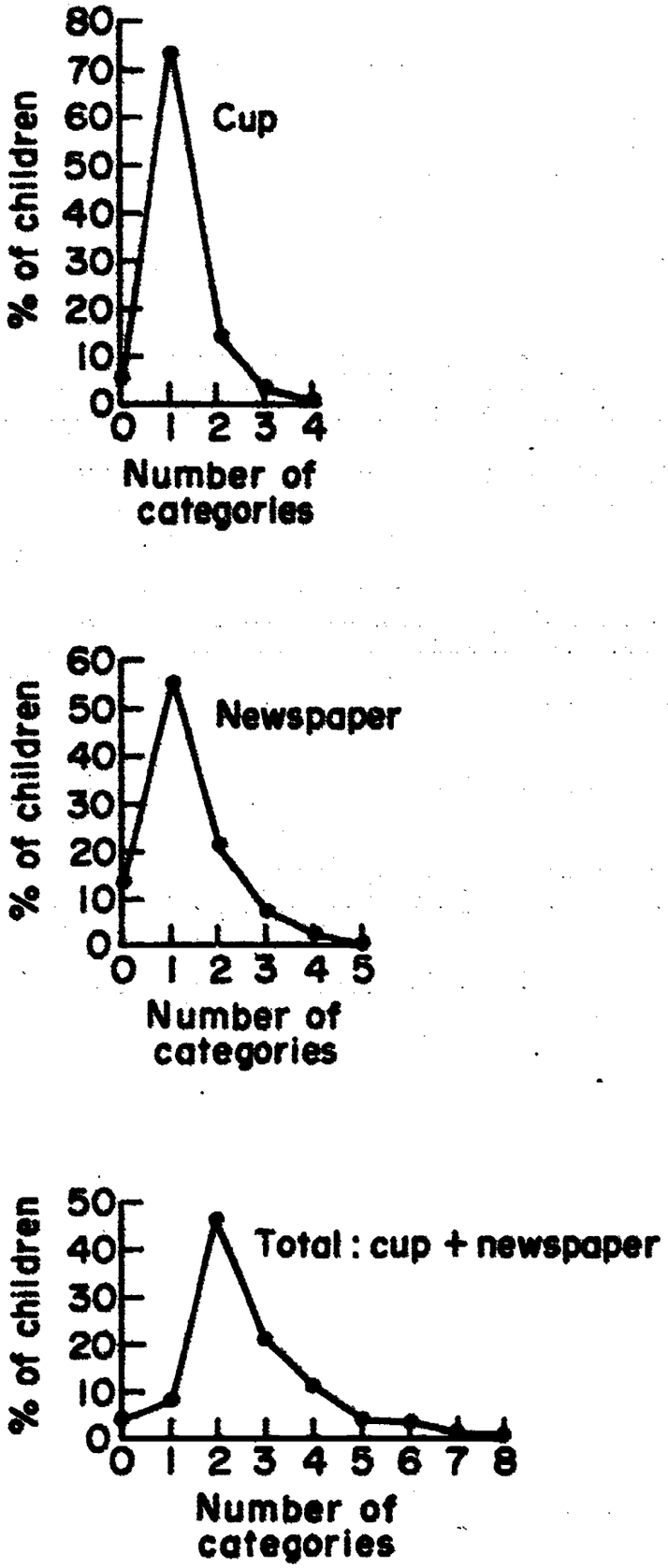


Figure 9. Distribution of category frequency for cup, newspaper and their combined total.

TABLE 19

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR

RESPONSE FREQUENCY IN UNUSUAL USES

Group	N or n	$\bar{X}$	SD	p
Cup				
All children	266	2.7	2.0	
Age: 3	62	1.7	1.5	
4	134	2.7	1.9	<.005
5	70	3.5	2.3	
Sex: Male	137	2.7	2.2	
Female	129	2.6	1.9	ns
Race: White	71	3.4	1.9	
Black	80	3.1	2.1	ns
Newspaper				
All children	266	1.6	1.2	
Age: 3	62	1.2	1.0	
4	134	1.5	1.2	<.005
5	70	2.1	1.3	
Sex: Male	137	1.6	1.4	
Female	129	1.6	1.1	ns
Race: White	71	2.0	1.3	
Black	80	1.9	1.5	ns
Total of cup and newspaper				
All children	266	4.3	2.9	
Age: 3	62	2.8	2.3	
4	134	4.2	2.7	<.005
5	70	5.6	3.1	
Sex: Male	137	4.3	3.1	
Female	129	4.2	2.6	ns
Race: White	71	5.4	2.8	
Black	80	5.0	3.2	ns

TABLE 20

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR

CATEGORY FREQUENCY IN UNUSUAL USES

Group	N or n	$\bar{X}$	SD	p
Cup				
All children	266	1.2	.7	
Age: 3	62	1.0	.7	
4	134	1.2	.6	ns
5	70	1.3	.7	
Sex: Male	137	1.2	.6	ns
Female	129	1.2	.7	
Race: White	71	1.4	.8	ns
Black	80	1.3	.7	
Newspaper				
All children	266	1.4	.9	
Age: 3	62	1.1	.8	
4	134	1.4	1.0	<.005
5	70	1.6	.9	
Sex: Male	137	1.3	1.0	ns
Female	129	1.4	.9	
Race: White	71	1.6	1.0	ns
Black	80	1.6	1.1	
Total of cup and newspaper				
All children	266	2.5	1.3	
Age: 3	62	2.1	1.4	
4	134	2.5	1.4	<.005
5	70	2.9	1.2	
Sex: Male	137	2.5	1.4	ns
Female	129	2.6	1.3	
Race: White	71	2.9	1.5	ns
Black	80	2.8	1.6	

(see Vol. I, pp. 25-28), all the means are considerably lower than those previously obtained with children of nursery school and kindergarten age. Two possibilities come readily to mind to aid in understanding these differences. First, Unusual Uses was administered second from last in the Inventory. Therefore, fatigue may be a factor in the lower scores. Second, the previous investigators presented the child with an actual cup and newspaper, whereas the psychometrists of this study did not. The presence of the actual items may have stimulated response production. Until these possibilities are further explored, the differences obtained between this and the two prior studies, which utilized Unusual Uses with children similar in age, cannot be interpreted.

Relationship of Scores to Selected Ratings of the Child and to Dog and Bone Scores

As in the case of Dog and Bone, Unusual Uses scores (total frequency and total category frequency) were analyzed for their relationship to selected ratings of the child made by his teacher and by the examiner. The same ratings and methods of analyses were involved.

Unusual Uses scores bore no significant relationship to any of the single-item ratings by the teacher, neither those which might be expected to show a relationship (uses imagination, overcomes obstacles, is curious) nor those assumed to be unrelated (follows directions, follows rules, likes to have help). There were likewise no relationships between the two Unusual Uses scores and teacher ratings of task orientation (.15 and .09) and distractibility (-.10 and -.09). The relationships between Unusual Uses and examiner ratings, however, were significant and positive, although by no means very large. There were correlations of .24 and .22 with social competence and correlations of .28 and .26 with task competence.

It will be recalled that correlations of total response frequency in Unusual Uses and the two Dog and Bone scores, discussed in Chapter 3, were .21 and .24. These correlations reflect only modest relationship between the two measures. These data were interpreted as indicating that, although the two measures purport to assess "creativity," they might very well measure quite different aspects of creativity. The additional findings reported in this and the preceeding section suggest that additional possibilities should be considered. Unusual Uses, at least as it appears in this inventory, may also be measuring fatigue factors and general competence in the assessment situation. Further analyses should explore the relationship between Unusual Uses and measures of verbal ability and verbal fluency since it may also relate to these variables.

Summary and Recommendation

Unusual Uses can be administered with ease and scored reliably (except for originality). The relationship of scores with age and the lack of relationship with sex and race suggest that this component assesses a cognitive skill. However, further work should explore more precisely the issues of what this component measures and how its results are influenced by its placement in the sequence of Inventory components.

CHAPTER 5

THE MANY-SPLENDORED CUBE

The Cube was designed to assess exploratory behavior and the range of sensory modes employed by the child in exploring his environment. It consists of a 4-inch hollow plexiglass cube with various visual, tactual, olfactory, and auditory properties. (It is pictured in Appendix C of Volume I.)

Instructions direct the psychometrist to present the cube to the child while saying, "Here is something for you to play with while I work on these papers for a minute." The psychometrist then unobtrusively observes the child and notes the behaviors he engages in. Ratings are made of the child's spontaneous initiation of activity with the cube and his involvement with the task.

Administration

The psychometrists reported that in general both they and the children had enjoyed the item very much. Their specific comments regarding experiences with the cube focused upon a) the wide variety of behaviors engaged in by the children, b) their concern that some active and/or curious children might break, rip, or tear apart the cube, and c) difficulty with the rating of "listens to cube." With regard to the latter point, it was felt that many children indeed seemed to be attending to the auditory properties of the cube even though they did not bring the cube close to their ear in order to do so.

Scoring

The following items were scored and tallied for purposes of analysis:

Task initiation: spontaneous or not

Task involvement: a 4-point rating from hesitant, hardly interested (1) to eager, enthusiastic exploration (4)

Comments made by child: none; 1-2; 3 or more

Questions asked by child: none; 1-2; 3 or more

Specific question ("What's inside the box?") asked by child: present or not

Comments and questions related to certain features of the cube:

- pictures (stickers)
- colors
- texture/appearance
- sound
- smell
- sum of the above tallies

Manipulative and exploratory behaviors: presence of behaviors in each of the following categories:

- turns cube over
- shakes, and/or listens; brings cube to ear and shakes
- "peels" (or attempts to peel)
- drops
- pounds ... taps cube on table
- rubs on table
- tries to open or pull apart the cube
- feels surfaces
- feels edges, corners
- rubs against cheek, etc.
- "tastes" cube (puts in mouth)
- smells, sniffs cube
- other behavior not able to be categorized in above categories
- unique behavior; mentioned less than once per 100 records

Interrater agreement in categorizing the childrens' responses, determined for each category, ranged from 97% to 100%. These findings give excellent confidence in the reliability of scoring.

Findings

The results reported here are based on the 278 children who were given this component of the Day Care Inventory.

Initiation and Involvement

Ninety-six percent of the total sample of children were rated as spontaneously initiating activity with the cube. There were no significant age or sex differences in this rating. A significant race difference appeared, however. All of the black children (vs. 92.6% of the white children) were rated as showing spontaneous initiation (χ^2 = 4.16, df = 1, p = .04).

The distribution of ratings for task involvement with the cube is presented in Figure 10. Roughly three-fourths of the children (77.4%) were rated on the "eager, enthusiastic exploration" half of the scale. There were no age- or sex-related differences in the involvement

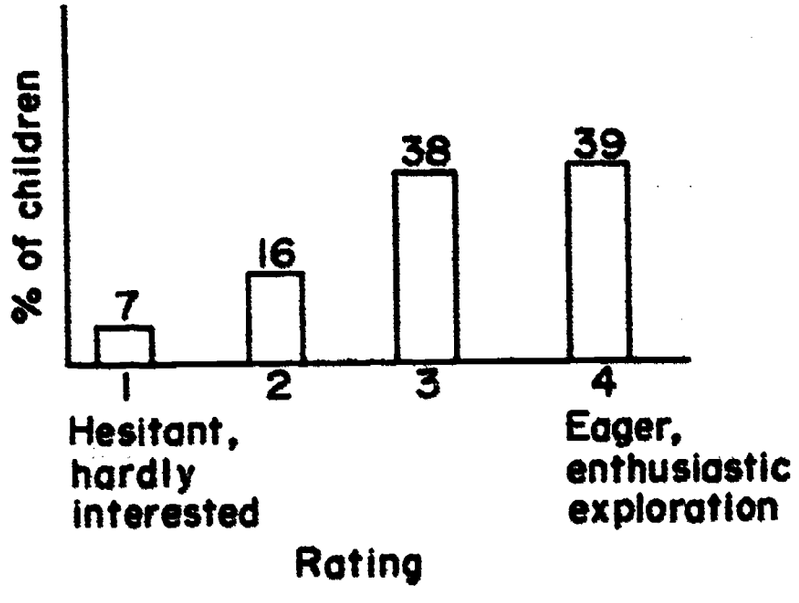


Figure 10. Task involvement with the cube.

ratings, but race was related to differences in ratings. The distribution of ratings for black children showed more ratings in the "eager, enthusiastic exploration" categories than did the distribution of ratings for white children ($\chi^2 = 8.61$, $df = 3$, $p = .04$).

In summary, these initiation and involvement ratings suggest that the cube is intriguing to children and prompts their active exploration. One might have expected some sex and age differences, but none appeared. Black children seemed more actively involved and showed greater spontaneous initiation of activity with this task than white children.

Verbalizations Made by the Children

Since the level of interest in the cube seemed high, the comments and questions about the cube were examined. Results indicate that comments were made by about half (52%) of the children and questions were asked by essentially the same proportion (51%). A chi-square analysis indicated that a child was more likely to have engaged in both verbal behaviors or neither than he was to have engaged in just one. Thirty-one percent both asked questions and made comments, 34% did neither, 18% asked questions only, and 17% made comments only ($\chi^2 = 22.94$, $df = 1$, $p < .00001$).

There were no statistically significant age, sex, or race differences in percentages of children asking questions or making comments about the cube. Racial comparisons, while not significant at the .05 level of significance, nevertheless seem interesting to note. White children seemed more likely than black children (51% vs. 39%) to make comments ($\chi^2 = 2.01$, $df = 1$, $p = .16$), while black children

seemed more likely than white children (56% vs. 43%) to ask questions ($\chi^2 = 2.30$, $\underline{df} = 1$, $\underline{p} = .13$).

Certain cube characteristics referred to in the children's questions and comments were separately analyzed. Figure 11 presents the percentage of children verbalizing about the following aspects of the cube: the contents of the cube (e.g., "What's inside it?"), the sound produced (e.g., "It rattles"), surface colors or other visual and tactual properties (e.g., "This is shiny," "This is bumpy"), and the scent (after shave lotion) emanating from the cube.

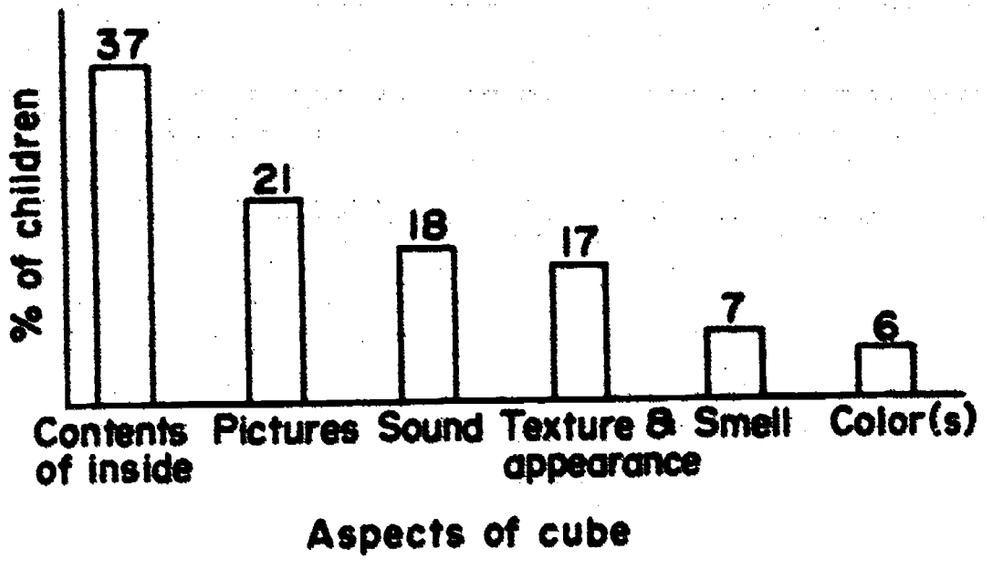


Figure 11. Aspects of the cube upon which children commented and asked questions.

It can be seen from Figure 11 that over one-third of the children commented or questioned about the contents of the cube, an aspect

which was not definitely apparent from the directly perceivable properties of the cube. (The cube contained a small bell and a paperclip; together they made a rattle which was highly ambiguous.) To this third could also be added the 18% who commented or questioned the sound produced but who did not specifically ask, "What's inside the box?" Other properties of the cube (color, textures, scent) easily and directly known by observation, were less frequently the focus of comments and questions. These data suggest that the unknown or ambiguous qualities of the cube more frequently stimulated the child's verbalizing than its obvious sensory properties, attractive as they were designed to be. These data suggest, but do not test, the hypothesis that stimulating a child's curiosity by ambiguous or incomplete input is more effective in stimulating the verbalization of young children than any single, attractive, but informationally, complete type of input.

It is unfortunate that a strict comparison of verbalizations related to the various sense modalities could not be made with these data. First of all, the modality utilized by the child could not be ascertained in some cases. In the texture and appearance category, touch and vision could not be separated as a basis of the child's verbalization. One could not tell, for example, whether a child based the comment, "It is soft," on tactual or visual cues. In a similar vein, the child who asked about the contents of the box may have been prompted by auditory cues or logic based on visual cues (i.e., This looks like a box; boxes have contents; therefore something may be inside). When examined in the light of these limitations, the

data summarized in Figure 11 indicate that the cube stimulates verbalizations related to its visual and auditory properties primarily and that it stimulates verbalizations related to touch and smell much less frequently.

There were no significant age differences in any category of stimulus-related verbalization, with the exception of that for verbal references to the pictures on one surface of the cube. In that category, more 5-year-olds than 3- or 4-year-olds made comments or asked questions.

No significant sex differences appeared for verbalization in any of the stimulus-related categories scored.

One significant race difference appeared in stimulus-related verbalization. More white children (21%) than black children (8%) verbalized about texture and other surface qualities ($\chi^2 = 4.14$, $df = 1$, $p = .04$). A trend toward significance ($p = .07$) occurred in the finding that black children (45%) more frequently than white children (30%) queried about the contents of the cube. This last finding is congruent with the teacher's Child Ratings of the black children as more curious than the white children.

Taken together, these findings regarding verbalizations related to various stimulus properties of the cube suggest that the verbalizations are <u>not</u> a function of language development per se. Otherwise more age and sex differences would have appeared. The paucity of significant findings in all subgroup comparisons, however, highlights the question of just what these verbalizations may be related to. Further analyses involving comparisons with other Inventory components

may lead to the answer of this question, but without positive evidence to the contrary it must be assumed that these particular verbalization scores, while telling us the relative "pull" of cube components, tell nothing understandable about the child.

Manipulative and Exploratory Behaviors with the Cube

The findings for the manipulative and exploratory behaviors of the child with the cube parallel those for verbalizations. The frequencies with which the children engaged in various behaviors with the cube can be described, but they do not appear related with any frequency or pattern to age, sex, or race.

The frequency with which the children engaged in various manipulative and exploratory behaviors with the cube are presented in Figure 12. It can be seen that general manipulative activities, such as turning the cube and feeling its surfaces and edges, were frequently occurring behaviors, as was the more specific activity of shaking the cube.

There were only two age-related differences among these behaviors. Shaking the cube was engaged in by significantly more 4-year-olds (87%) than either 3-year-olds (78%) or 5-year-olds (74%), which did not differ significantly from each other ($\chi^2 = 6.22$, $df = 2$, $p = .04$). Although dropping or attempting to drop the cube was an infrequent activity in general (2.5%), it was engaged in more frequently by 5-year-olds (6.8%) than by either 3-year-olds (1.6%) or 4-year-olds (.7%). While this finding is statistically significant ($\chi^2 = .02$, $df = 2$, $p = .02$), its meaningfulness appears close to trivial because of the small frequencies involved.

No sex differences in these behaviors were found and only one

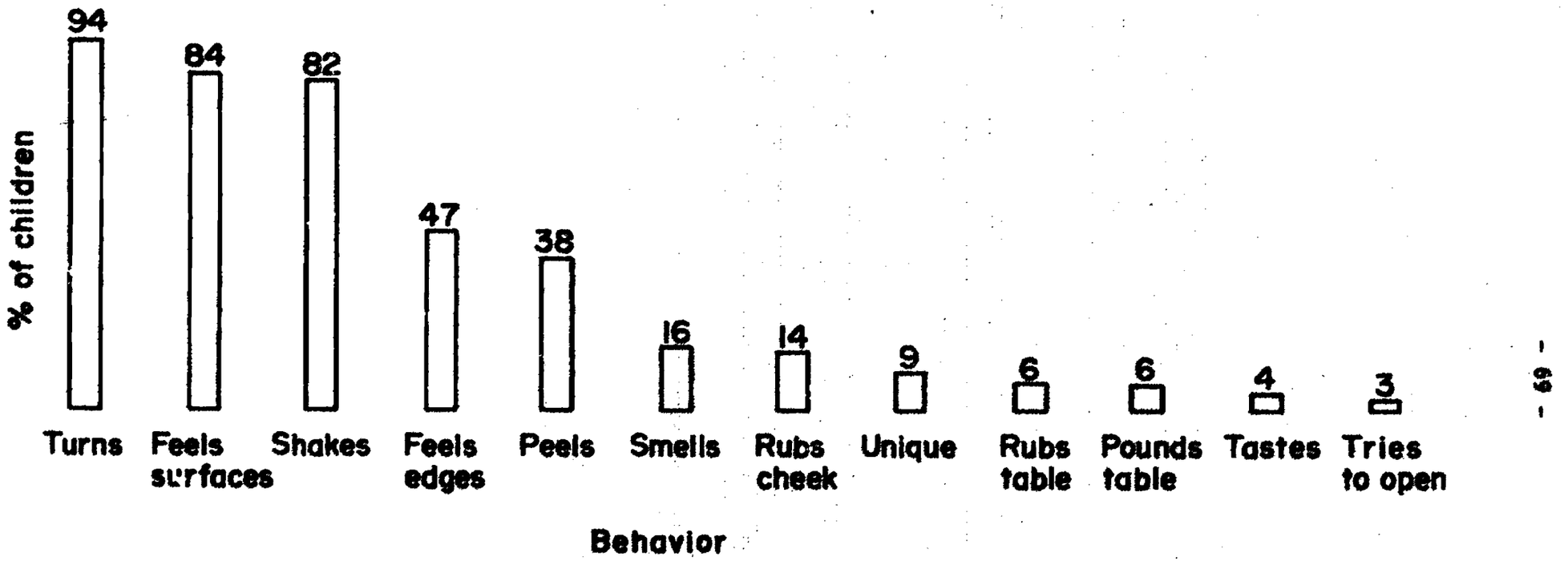


Figure 12. Manipulatory and exploratory behavior engaged in with the cube.

race-related difference appeared, that for peeling or attempting to peel a surface of the cube. More black children (57%) than white children (36%) engaged in this behavior ($\chi^2 = 5.69$, $df = 1$, $p = .05$).

A total score comprised of all behavior items was computed for each child. While the distribution of these scores showed excellent shape (normal) and range (0 to 9), there were no age, sex, or race relationships with this total score.

Summary and Recommendations

The Many-Splendored Cube is an interesting item for children. While it can be determined what the children do in response to the cube, the domain in which this component reflects any individual differences remains a mystery. The measure is unrelated in any systematic way to the three variables of age, sex, or race. If further analyses can determine that the cube represents an age-, sex-, or race-free measure of something with assessment value, a unique measure may have been developed. Otherwise the cube should be dropped from the Inventory.

CHAPTER 6

PICTURE NAMING

Picture Naming was modeled after traditional measures of a child's expressive language development. It assesses his store of nouns and his ability to use them to identify common objects. Materials for this component are 20 simple drawings presented one at a time. The examiner introduces the task by saying, "Now I want you to look at some pictures with me and tell me what the things are called." On presentation of each picture the examiner says, "What is this?" and/or "What do you call it?" Examiner instructions call for the use of reinforcing, encouraging comments to the child at specific points.

Administration and Scoring

The examiners were unanimous in their feeling that this component was pleasurable for children. Typical verbatim comments were the following: "This is a fun task for them...even if you don't get the research information you want, it is a rapport builder," "They know they can do it," "It is a successful experience for children, particularly little ones."

Scoring was not done by the examiner at the time of administration, as it ordinarily is in similar measures, but was done at a later time by other staff members. In scoring the children's responses, credit was given for synonyms (e.g., "brella" for umbrella) and childrenese (e.g., "horsie" for horse). The high reliability of scoring is indicated by the 99.6% agreement between two scorers who independently scored 25 cases (500 responses).

Findings

Distribution of Scores

This component was administered to 279 children. The total scores of this 20-item measure ranged from 0 to 19, with a mean total score of 14.6 and a standard deviation of 2.2. The distribution of total scores is further illuminated in Figure 13, where it can be seen that the scores are skewed toward the higher scores but are relatively symmetrical about their mean. The impression of the psychometrists that this is an easy component for children is borne out by these findings.

Item Analyses

The difficulty of each item and its discriminability are indicated in Table 21. Item difficulty is reflected in the percentage of children passing the item. Item discriminability is represented by a biserial correlation between success or failure on an item and the total scores achieved for the component. It is also reflected by the average total scores of children who pass the item and of children who fail the item.

First, the findings in Table 21 again point out that this is indeed an "easy" component. Eleven of the 20 items were correctly named by over 90% of the children. The difficulty of the remaining nine items was more evenly distributed, but still was skewed toward ease of passing. From the standpoint of desirable spread of item difficulty, this component is very weak.

With regard to how well passing or failing each item relates to total score on the component, it is noted that the biserial correlations

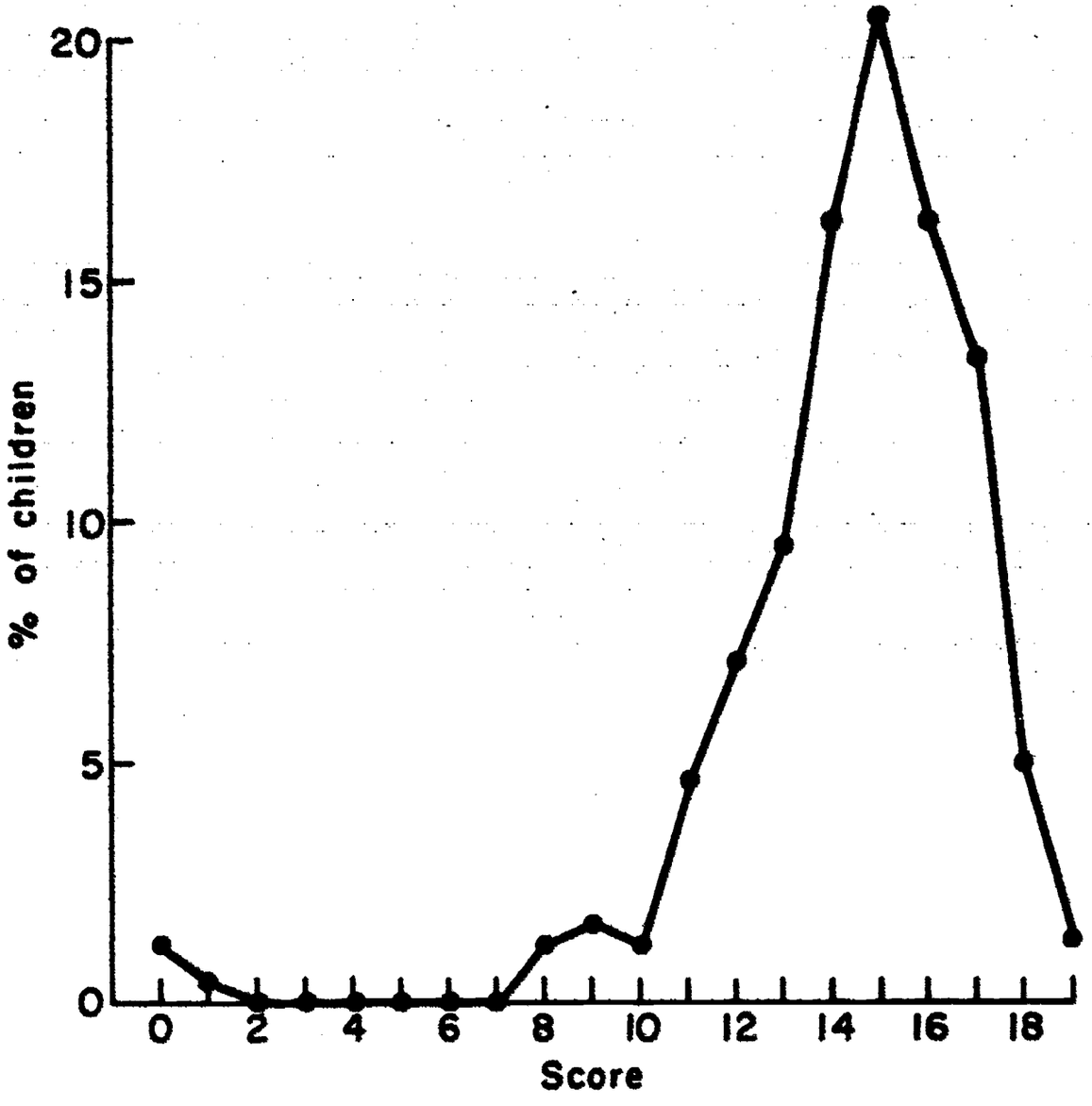


Figure 13. Distribution of the total scores for the Picture Naming component.

TABLE 21

ITEM ANALYSES OF PICTURE NAMING

Item (in order of increasing difficulty)	% of children passing item	Biserial correlation with total score	Average total score of children who pass item	Average total score of children who fail item
Tree	98	.46	14.6	7.2
Knife	97	.57	14.7	9.2
Fork	96	.59	14.7	9.4
Telephone	96	.71	14.7	8.6
Boat	95	.69	14.8	9.4
Purse	95	.68	14.8	9.8
Hand	95	.53	14.7	10.5
Key	95	.55	14.7	10.6
Hat	94	.76	14.8	9.5
Horse	93	.51	14.7	11.2
Tire	93	.51	14.8	11.3
Umbrella	89	.69	14.9	11.2
Guitar	71	.60	15.2	12.7
Football	66	.60	15.3	12.9
Coat	65	.34	15.0	13.6
Leaf	58	.75	15.7	12.8
Spider web	35	.75	16.4	13.5
Fish hook	14	.60	17.0	14.1
Pitcher	8	.43	16.7	14.3
Door hinge	1	.27	17.8	14.4

were all positive. Averaging about .58, they are acceptable but not as high as would be desired. The differences in total scores of children who passed and children who failed the items were all significant. All the items were successful in "discriminating" but were not as excellent in doing so as one would wish.

The psychometric properties of this component could be improved by removing some easy and less discriminating items and by adding more difficult and more highly discriminating ones. The point that easy items are happy experiences for children should be taken seriously, however. It may be best to keep the easy items as "fillers" for rapport's sake and to omit scoring them for measurement's sake. At the same time items could be added to improve the difficulty range and discriminability level.

The interitem consistency of this component was .68, as assessed by the Kuder-Richardson formula 20. Moderate homogeneity is thus indicated.

Subgroup Analyses of the Total Scores

Presented in Table 22 are findings related to possible subgroup differences. It can be seen that only the age differences were significant, as they would be expected to be inasmuch as language development is strongly related to age. These age differences were significant overall ($F = 46.63$, $df = 2/275$) and between each age group ($3 < 4 < 5$, by Tukey's test).

No sex or race differences were found. Inasmuch as some vocabulary tests are felt to be culturally biased in favor of white children, it is encouraging to note that this measure appears "culture-fair" in this regard.

TABLE 22

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR

THE PICTURE NAMING COMPONENT

Group	N or n	$\bar{X}$	SD	p
All children	278	14.6	2.2	
Age: 3	63	13.0	2.5	
4	141	14.5	1.8	<.01
5	74	16.1	1.6	
Sex: Male	142	14.9	2.3	ns
Female	136	14.3	2.0	
Race: White	73	15.2	2.0	ns
Black	84	14.7	1.9	

Recommendations

This component should be retained in the Inventory. It is a "hit" with the children, and it can be administered easily and scored reliably. Its goodness as a measure should be improved by increasing the number of more difficult and more highly differentiating items.

CHAPTER 7

PICTURE INTERPRETATION

The Picture Interpretation component was included in the Inventory because it might be of use, as similar items traditionally have been, as a rough indicator of a child's language development. It was further felt that, inasmuch as a group care environment was portrayed, the content of the child's response might shed light on aspects of the child's perception of that environment. (The picture is shown in Appendix C of Volume I.)

The children were told "Look at this picture and tell me all about it." A second request, "Tell me all about this picture," was made if a child did not respond after 10-15 seconds, and the examiner stated, "Tell me more about the picture," if the child had responded by enumerating only one or two items in the picture. At the completion of this modified free response phase, the examiner asked four questions designed to tap the child's perception of (a) what the group of adult and children were doing, (b) what the solitary child was doing, (c) how the solitary child was feeling, and (d) what sort of place was depicted.

Administration and Scoring

Examiners' feedback indicated that administration of this component was straightforward. Expectedly, the verbatim recording of lengthy responses given at a fast rate was reported to be difficult. The psychometrists also concentrated upon the frequency with which the question regarding the solitary child, "How does this child feel?," was interpreted concretely, leading to such responses as "With his

fingers." They were the first to suggest, as does the analysis of this question in the following pages, that a rewording of this question is in order if such concrete interpretations are to be avoided.

Responses were tabulated to indicate the child's type of response (enumeration, description or interpretation), the number and identity of items mentioned, and the possible affect spontaneously associated with the picture or the people depicted. In regard to the specific question asked about the group pictured, aspects tabulated were: focus of response (children, adult or neither), the activity specified, and any spontaneous identification of the adult in the group (unspecified female, teacher or mother). The specified activity and feeling of the child were tabulated regarding the solitary child. Finally, the child's identification of the place depicted was tabulated. Categories for tabulating various aspects of Picture Interpretation which have not been mentioned in this section are apparent in the following presentation of the results.

Reliabilities of the tabulations were assessed by the independent scoring of 40 protocols by two persons. Reliabilities in all categories equalled or exceeded 90% agreement.

Findings

Type of Response

Of the 275 children given this component, 90% were able to give scorable responses. There were no age, sex, or race differences in the ability to give a scorable response. When type of response was classified into (a) simple enumeration of depicted items, (b) description of the picture, and (c) thematic interpretation of the picture,

31% of the children utilized simple enumeration, 22% utilized description, and 47% utilized interpretation. There were no sex or race differences in the distribution of responses so classified. There was a trend, however, toward significant age differences in the type of response. Enumeration declined with age, and interpretation increased with age ($\chi^2 = 4.68$, $\underline{df} = 4$, $\underline{p} = .10$). When enumeration responses are compared with description and interpretation responses combined, there are again no sex or race differences, but the age differences become highly significant ($\chi^2 = 7.48$, $\underline{df} = 2$, $\underline{p} = .03$). Table 23 presents the distribution of nonscorable, enumeration, and description/interpretation responses according to the age of the child. These age-related findings suggest that the Inventory picture may be utilized, as others traditionally have been, as an index of language development. The poorer age differentiation obtained when description and interpretation are not combined supports the maintenance of the enumeration vs. description/interpretation scoring for future work.

TABLE 23

TYPE OF RESPONSE GIVEN TO PICTURE INTERPRETATION

Age of child	Type of response		
	Nonscorable	Enumeration	Description/interpretation
3 years[a]	9.4%	37.5%	53.1%
4 years[b]	9.4%	28.8%	61.9%
5 years[c]	11.1%	16.7%	72.2%

[a] n = 64
[b] n = 139
[c] n = 72

Stimulus Characteristics of the Picture Apparent in the Children's Free Response

Findings which describe the "stimulus pull" of the particular picture utilized are presented next. These findings are based upon the frequency with which various human and nonhuman categories were mentioned in the child's free response to the picture. These are purely descriptive data, and comparative data from other groups of children, particularly those in other forms of group care or in no group care, should be sought.

Table 24 presents the percentage of children whose responses indicated attention to the different social groupings protrayed in the picture, together with the few findings of subgroup differences in these percentages. It can be seen from Table 24 that the four social groupings were mentioned with approximately equal frequency; these groupings, despite their difference in position and content, appear to have about the same stimulus pull. No subgroup differences were found which have

TABLE 24

PICTURE INTERPRETATION RESPONSES TO SOCIAL GROUPINGS

Social grouping	Percent of children responding[a]	Subgroup differences
Solitary child	34	
Child/adult pair: child on adult's lap	40	age: 3<4<5, p = .15
Child/child pair: two children at table	37	age: 3 = 4<5, p = .06
Group of 1 adult and 3 children	44	race: white < black, p = .06

[a] N = 248

robust statistical significance, but one race-related trend was found (white children mention the adult/child group more frequently than black children) and two age-related trends were found (mention of adult/child pair increases with age and mention of the child/child pair is less frequent for 3-year-olds than for 4- or 5-year-olds.

Of the various human and nonhuman categories tallied, the average total number mentioned by the children was 4.2 (SD = 2.7). This average was significantly related to age. Three-year-olds mentioned an average of 3.1 items, 4-year-olds averaged 4.4 items, and 5-year-olds averaged 4.8 ($F = 7.22$, $df = 2/245$, $p = <.01$). Approximately equal tallies were obtained for human and for nonhuman references: an average of 2.0 (SD = 1.2) references to human and an average of 2.2 (SD = 2.3) references to nonhuman objects. Both human and nonhuman references were related to age. For nonhuman objects, the average increased with age: 1.5 at age 3, 2.1 at age 4, 2.8 at age 5 ($F = 5.29$, $df = 2/245$, $p = <.025$), while for humans 3-year-olds gave fewer references (1.6) than either 4- or 5-year-olds (2.2 and 2.0, respectively) ($F = 6.18$, $df = 2/245$, $p = <.025$). No sex or race differences were found in these measures.

Table 25 represents the "stimulus pull" of various aspects of items in the picture. In some instances the exact stimulus referred to is unclear. "Child," for example, can refer to any of the children pictured, and it is not certain that "mommy" always refers to the adult holding a child in her lap. The exact stimuli involved is a clearer matter in the case of nonhuman objects.

TABLE 25

"STIMULUS PULL" OF ITEMS MENTIONED IN FREE RESPONSE TO PICTURE INTERPRETATION

Item	Percent of children mentioning[a]	Subgroup differences
Humans		
Child, sex unspecified	56.0	
Child or children, unspecified	48.8	age: 3<4 = 5, p = .0005
Child, named	1.6	
Baby	19.0	
Adult female, e.g., "lady"	10.9	
Teacher	21.4	
Mother, mommy	41.9	age: 3<4 = 5, p = .05; sex: girls>boys, p = .01
Grandmother	4.0	
Daddy	4.8	
Nonhuman objects		
Fish	39.7	age: 3<4<5, p = .005; sex: girls>boys, p = .04
Doll	35.9	sex: girls>boys, p = .09 race: white<black, p = .06
Clock	22.6	age: 3<4<5, p = .001
Picture(s)	22.2	age: 3<4<5, p = .001
"Toys"	20.6	
Furniture (e.g., chair, table)	19.0	
Waste basket	16.1	age: 3<4 = 5, p = .04
Car	12.9	sex: boys>girls, p = .10
Book(s)	12.5	
Blocks	8.1	
Architecture (e.g., door, shelves)	6.9	age: 3 = 4<5, p = .004
Identity of place		
School	10.9	age: 3 = 4<5, p = .005
Other	1.6	

[a] N = 248

Regarding human references, it can be seen in Table 25 that responses involving the children pictured were somewhat higher than those involving adults, as might be expected from the greater number of children portrayed. The age-related findings regarding the humans show that 3-year-olds make fewer nonspecific child references (e.g., "these children," "they") than do 4- and 5-year-olds. This finding might be related to the older children's tendency to group percepts, to generalize, and to describe rather than to enumerate specific children. The finding that significantly more 4- and 5-year-olds than 3-year-olds mention "mother" may be hypothesized to be a function of their greater role differentiation of mothers and teachers (i.e., mothers hold children on their laps). A similar explanation may underlie the significantly greater mention of "mothers" by girls than by boys.

The fish and the doll were the most "popular" nonhuman objects in terms of frequency of mention. Interestingly, they are the two objects which are the least inanimate. The age differences found are assumed to reflect the developing child's increasing attentional-perceptual span, his ability to notice more in his environment, and his increasing ability to incorporate more into his verbal response.

The sex differences found in both the human and nonhuman categories (girls more frequently mentioning doll, fish, and mommy and less frequently mentioning car) is hypothesized to reflect the influence of differential sex role experiences and expectations on the selective perception and/or verbal response of young children.

A final comment regarding the stimulus pull of the picture as reflected in the child's free response concerns the affect or emotion

associated with the picture. Spontaneous expressions of affect or evaluation regarding either the picture as a whole (e.g., "That's a pretty picture") or persons depicted (e.g., "He's real mad") were extremely infrequent. Only seven children made such responses. The picture, thus, is seen to have no strong affective connotations.

Stimulus Characteristics of the Adult-Children Group Apparent in Response to Direct Questioning

If a child had not spontaneously referred to the group in his free response, the examiner pointed to the group and asked, "What are they doing?" Table 26 presents the answers to this question given by those 224 children who specified a particular activity. The most frequent responses were rather bland, nonspecific, global references to sitting, talking, and listening, although a good proportion of responses delineated the activity more precisely as that of an educational nature. Much less frequently the activity was seen as play or as an occasion during which the teacher was imparting directions and discipline.

Whether the focus of the child's response to this grouping was on the adult, the children, or neither in particular was also of interest. Twenty-two percent of the children's responses focused upon the adult, 24% focused upon the children, and 54% had neither focus. There were no significant age differences in this variable, but sex and race differences were found. Boys and girls focused upon the adult to the same extent (22%), but girls focused upon the children more than boys (30% vs. 18%), while the remaining category of no particular focus was more frequent for boys than girls (60% vs. 48%) ($\chi^2 = 5.45$, $df = 2$, $p = .07$). The salient finding here is the greater focus of the girls

upon the children in the picture. The prominent finding regarding race comparisons is the greater focus on the adult by white children (32% vs. 13% for black children). There was a complementary difference in the category of no particular focus (38% for white children vs. 58% for black children) and no difference in the frequency of focus upon the children depicted (30% vs. 29%, respectively).

TABLE 26

PERCEIVED ACTIVITY OF THE GROUP

Activity	Percent of children mentioning[a]
Sitting	22.3
Listening, talking	19.2
Educational: teacher is teaching or children are learning	17.4
Educational: singing, storytelling	13.8
Playing, not further elaborated	10.7
Direction and discipline: teacher telling children what to do	3.6
Other	12.9

[a] N = 224

Whether or not she was the focus of the child's response to the question about the group, the adult woman in the group was mentioned by 95, or about 40%, of the children. Whether she was described as a teacher, a mother, or given an indeterminate role ("her," she," "lady")

was noted. While more than half of the children (57.9%) identified her as a teacher, about one-third (31.6%) perceived her as a mother, and a small percentage of the children (10.5%) gave indeterminate responses. The perception of the role of the adult woman did not significantly vary with the child's age, sex, or race. Thus, while to adult eyes the adult may be perceived as a teacher, young children do not necessarily share this perception.

Stimulus Characteristics of the Solitary Child Apparent in Response to Direct Questioning

Two specific questions referring to the solitary child in the picture were asked by the examiner if the information desired was not already available in the child's free response. These questions were, "What is this child doing?" and "How does this child feel?" Eighty-five percent of the children gave scorable replies to the first question, and 74% gave scorable replies to the second. There were no age, sex, or race differences in these percentages.

Table 27 presents the percentage of children mentioning various activities ascribed to the solitary child. The solitary child was rather routinely seen as engaged in some sort of play. Overall chi-square tests were performed to detect any sex, age, or race differences in perception of activity, significant results obtaining for age and race (for age, $\chi^2 = 34.12$, $df = 18$, $p = .01$; for race, $\chi^2 = 18.70$, $df = 9$, $p = .03$). Age changes appeared most clearly in a decrease in the "other" responses, in an increase in the specific playing responses, and in an increase in the playing with playdough responses. Race differences indicated that white children, when compared to black

children, gave more specific playing responses, more playdough resonses, and fewer sitting and "other" responses.

TABLE 27

PERCEIVED ACTIVITY OF THE SOLITARY CHILD

Activity	Percent of children mentioning[a]
Playing: specific activity	30.1
Playing: not further elaborated	19.2
Playing with playdough	14.8
Sitting	9.2
Looking	5.7
Spilled something	3.5
Crying	1.7
Other	15.8

[a] N = 229

The feelings ascribed to the child are presented in Table 28. The results in Table 28 suggest that the solitary child is appropriately ambiguous in terms of what his feelings might be since there was fairly good distribution of the responses over positive and negative feelings. About one-third of the children perceived the solitary child as having positive (good, happy) feelings, and about two-fifths perceived negative (bad, sad, sick) feelings. Inasmuch as it was the solitary child's perceived emotional feelings which were to be assessed, it is disconcerting to discover the relatively high percentage of children (18%) who interpreted the word "feel" in our question, "How does this child feel?," as referring concretely to tactual feelings and replied, e.g.,

"With his hands." A rephrasing of the question is thus in order.

TABLE 28

PERCEIVED FEELING OF THE SOLITARY CHILD

Feeling	Percent of children mentioning[a]
Positive: good, happy	31.8
Negative: bad	15.9
Negative: sad	15.4
Negative: sick	10.0
Tactual feeling, e.g., "With his hands"	17.9
Other	9.0

[a] N = 201

Overall chi-square tests were also performed on these data regarding the feelings of the solitary child to assess possible age, sex, or race differences in this perception. Although there was a decline with age in the perception of positive feelings (42% at age 3, to 20% at age 5) and an increase in bad and sad feelings, the overall chi-square did not reach statistical significance. There were no significant sex differences but the race comparison was highly significant ($\chi^2 = 15.11$, df = 6, $p = .02$). Black children more frequently perceived positive feelings than white children (44% vs. 20%, respectively), and black children less frequently interpreted the question as referring to tactual feeling (8% vs. 30%, respectively).

Identification of Setting

Although to adult eyes the scene of the picture might be quite

readily identified as a room in a day care center or nursery school, it is an empirical question whether indeed young children would so perceive the setting. Accordingly, examiners asked each child, "What sort of place do you think this is?" The first finding of interest is that only 76.6% of the children could give a scorable reply to this question. Significantly more girls (83%) than boys (71%) gave a scorable reply (χ^2 = 4.96, df = 1, p = .03). Age was also significantly related to the frequency of scorable replies: 67% of the 3-year-olds, 76% of the 4-year-olds, and 86% of the 5-year-olds gave scorable responses (χ^2 = 7.12, df = 2, p = .03). There were no significant race differences in this variable.

Table 29 shows how the children (those able to give scorable replies) identified the setting. Only about one-half of the children perceived the setting to be a school or day care center. About a quarter of the children perceived the setting as "a house." The remainder perceived it as something else or as a specific part of a house (e.g., doctor's office, kitchen). Thus the stimulus picture does not necessarily portray a day care setting to children.

There were no significant age or race differences in the perception of the setting, but age was significantly related to this perception (χ^2 = 13.21, df = 6, p = .04). Description of the setting as a house or other setting declined with age, while identification of it as a school or day care center increased with age (43% at age 3, 48% at age 4, 77% at age 5).

TABLE 29

IDENTIFICATION OF SETTING OF STIMULUS PICTURE

Setting	Percent of children mentioning[a]
School or day care center	55.2
House	24.8
Other	20.0

[a] N = 210

Summary and Recommendations

Picture Interpretation should be retained as a rough index of a child's language development. The findings that the picture is not routinely identified as a day care center and that the "teacher" is not always perceived as a teacher suggests that if the picture is to be used to explore a child's reactions and feelings about his day care experiences, the procedure or stimulus should be changed. One possible procedure would be to identify the scene for the child (e.g., "This is a place like your center and this is the teacher") and then ask a focused question regarding the matter of interest (e.g., "What is she telling the children?"). Further, in order to gain knowledge of the childs' perception of the emotional feeling of the solitary child in the picture, the question used will require rephrasing to avoid a child's interpreting the word "feel" in the tactual sense.

CHAPTER 8

GROSS MOTOR SKILLS

The Gross Motor Skills component represents an attempt to provide the Inventory with a rough index of a child's gross motor ability. It is comprised of four activities which have been utilized in previous tests and inventories: balance on one foot, hop in place on one foot, walk forward heel-to-toe, and walk backward heel-to-toe. The examiner demonstrates each item and records either time (for balance) or number of responses (for hop and heel-to-toe walks).

Summary of Findings and Presentation of a Revised Gross Motor Skills Component

Gross Motor Skills involved many difficulties and variations in administration. Analyses of results could not be undertaken for any purpose other than to provide the most global judgments regarding the improvement of the component. For this reason results for the Gross Motor Skills component (GMSI) are not discussed in the same manner as those for other components but are mentioned only when germane to the following presentation of a revised Gross Motor Skills component (GMSII), which is currently being pilot tested and which hopefully removes sources of difficulty and error in its precursor.

Balance

GMSI allowed balance on either foot and did not specify the orientation which the examiner should adopt in demonstrating the item. Hence, one child might attempt to balance on a preferred foot,

another might attempt to copy the examiner on a nonpreferred foot, etc. Two trials were required and a third was required if the child could not balance for 10 seconds during one of the two trials. Examiners were not clear regarding when to terminate a trial. They also were not clear regarding whether a child should be allowed, for example, to hold his free foot with a hand or prop the free foot on the other knee. Results from GMSI indicate that the average balancing time on the first trial was approximately 7 seconds, with about 15% of the children achieving balancing times over 10 seconds. Balancing time increased with age. It was felt that 15 seconds would provide a meaningful cutoff point.

GMSII structure and instructions for examiners attempt to take into account both the shortcomings and results of GMSI. They are as follows:

Balance: 4 trials; 2 on each foot; get up and stand <u>beside</u> (not facing) child
Say: "Now let's do something different."
Balance on right foot, not holding on to anything.
Say: "Watch me; I am standing on one foot."
After balancing 10 seconds, stand on both feet about 10 seconds. Then balance on left foot.
Say: "I'm going to see if you can stand on one foot. Do it for as long as you can."
(Don't let child hold on to one foot while he balances on other)
Terminate any trial at 15 seconds.
Record time and foot (left or right) the child chooses.
Give second trial on foot chosen in first trial.
Say: "That was good! Try it again with the same foot."
Record time.
Say: "Now do it with your other foot. See how long you can stand on the other foot."
(Point to child's other foot if necessary)
Record time and give last trial.
Say: "Good. Now do it again with that foot."
Record time.

The examiner records the child's performance as follows:

Circle foot used			Record time in seconds
Trial 1.	R	L	Time _____
2.			Time _____
3.	R	L	Time _____
4.			Time _____

Heel-to-Toe Walk

The heel-to-toe item called for the child to walk forward in such a manner that the heel of the advancing foot touched the toe of the foot in place (or no more than 1 inch from it). Because of lack of clarity in GMSI instructions, some examiners terminated (as intended) at 10 steps of any sort, but others let the child continue (e.g., up to 20 steps) until he had achieved 10 correct steps out of the total. Obviously data from two such different administrations are not comparable. Further, examiners reported difficulty in teaching the task to some children. Indications from GMSI findings are that about 15 to 20% of the children who were given this task could not do even one step successfully, about one-quarter to one-third could complete 10 steps successfully, and a fairly low percentage of the children completed the intervening one to nine steps. It may be that this item will turn out to be more successful as a pass-fail item (one or more steps versus none) than as a count item (where number of steps is assumed to be meaningful). Performance on this item increased with age.

Instructions for heel-to-toe for GMSII are the following:

> Heel-to-toe walk, forward: 2 trials. Stand beside child and walk heel-to-toe, forward, for 5 steps.
> Say (while demonstrating): "See how I am walking. This foot touches this foot. Now you do it."

Encourage child (e.g., "Keep going") to complete 10 steps. Terminate at 10 steps of any sort or if the task is obviously uncomfortable for him.
Record the number of steps in which heel and toe are no farther than 1 inch apart (by examiner's estimation).
Give second trial.
Say: "Now do it again!" Encouragement, termination and recording same as for Trial 1.

The examiner then records for each trial the number of steps, out of the 10, which are less than 1 inch apart.

The backwards heel-to-toe walk has been omitted from GMSII because it proved to be too difficult an item for the majority of the children. About one-third of the children could not succeed in making one such step on the first trial of this item in GMSI, and about one-sixth of the total number of children were not administered a second trial, it being judged too frustrating for them in the opinion of the examiners.

Hop

The hop item called for the child to hop on one foot, the examiner first demonstrating by hopping in place. The child could use either foot, and a maximum of 10 hops was allowed. About one-third of the children achieved 10 hops, so there was obviously not enough "ceiling" on this item. GMSII instructions for the hop item specify an upper limit of 15 hops, and one trial on each foot. They are as follows:

Hop - 2 trials
Stand beside child. Say: "Watch me." Hop in place 5 times. Stop and then say: "Now let me watch you! You hop! Hop as many times as you can." (No holding on to furniture, etc.)
Terminate at 15 hops or if child puts second foot down.
Record number of hops, foot used, and whether in place (1 foot radius) or over a distance.
Give trial on other foot.
Say: "That was good! Now hop on your other foot as many times as you can."
Recording same as Trial 1.

Recording is made as follows:

Circle foot used			In place (P) or over distance (D)?		Number of hops
Trial 1	R	L	P	D	_____
	R	L	P	D	_____

Concluding Comment

The preceding material has presented revisions in the Gross Motor Skills component which hopefully will improve the accuracy and reliability of measurement. A larger issue, however, is whether such a small sample of motor activities accurately represents the large domain of motor skills and the child's physical development. Despite the arguments that can be made for assessing the "whole child" the present Inventory may be only paying lip service to this desideratum by including such a small number of activities to represent a child's physical development. If further day care research involves specific focus on the child's physical development, it would seem necessary to have a larger inventory of gross motor skills.

CHAPTER 9

VISUAL DISCRIMINATION

Visual Discrimination was included in the Inventory as a measure of a child's ability to discriminate on the basis of visual cues. There are two subcomponents. The first, Similarities, involves finding which picture among four alternatives is similar or identical to a standard. The second, Differences, requires finding which picture of four is different from the others. (Sample items are presented in Appendix C of Volume I.) Each subcomponent is preceeded by an example intended to familiarize the child with the task and has 10 items presented on separate picture cards.

Examiner instructions for Similarities are as follows:

> Present first card (example).
> Say: "Here are some pictures. See this one here?" (point to picture to left of double line) "Now look at these pictures here" (slide finger along the other 4 pictures on the card). "Find the one of these" (pointing along the 4 pictures) "that looks JUST LIKE this one." (pointing to picture on left).
> Do example with the child. Show him correct picture if he misses it.
> Say: "Good! You (we) found it. You (we) found the one that looks just like this one." (pointing)
> Present items 1 through 10 saying, "Find one just like this one. Point to it." for each item.

Examiner instructions for Differences are as follows:

> Say: "Here is a different book of pictures."
> Present first card (example) Say: "Find the one that looks DIFFERENT."
> "Find the one that is NOT THE SAME as the others. Point to it."
> Do example with the child. Show him the different one if he misses it.

Say: "That's it! Good! That's the one that ooks different."
Present items 1 through 10 saying, "Find the one that is DIFFERENT, the one that is NOT THE SAME. Point to it." for each item.

Administration and Scoring

The psychometrists reported that this component was the most difficult for them and for the children. It is apparent that modifications in this component should be made which take into account the several matters they raised.

They reported that it was difficult for children to understand their tasks, particularly that for Differences. Many children apparently "caught on" at a point well into the series of items. Others, never grasping what they were supposed to do, developed random or position responses (e.g., always choosing the left-most item). It is understood that the concept of "different" is an emergent one for preschool children, so difficulties with this task are to be expected. The examiners suggested, however, that the difficulties be reduced by giving several examples beforehand. It was also suggested that the administration of Similarities and Differences be temporally separated so that the "set" for Similarities ("Find one that looks just like this one") would not carry over into the Differences task ("Find the one that is not the same") and confuse the child.

Further suggestions related to errors which remained in the items due to the haste with which this particular component was prepared for field assessment. They are errors which are important but easy to correct: a) make sure supposedly identical items are exactly

that, so that children do not become misled by minor and unintentional variations and b) randomize the position of the correct responses throughout the series. As currently constructed, some positions are considerably more frequent than others. There are even instances in which the correct response is to be found in an identical position over a sequence of items. It is the child's visual discrimination that is important, not his ability to learn position responses.

During administration of the Inventory, the examiners recorded the child's choice (A, B, C, or D) to represent its position among the alternatives. Scoring for correctness of the choice and totaling correct choices was done at a later time by other staff members (with 99.8% accuracy as determined by two independent scorings of 50 records).

Findings

Distribution of Total Scores

The Similarities subcomponent was administered to 267 children; the Differences subcomponent to 248. Inasmuch as Differences immediately followed Similarities, the decline of 19 children in the two totals probably represents those children who refused further participation at this point and those for whom the examiner felt it was best not to proceed with the Differences subcomponent.

The total scores for both Similarities and Differences ranged from 0 to 10. The mean total score for Similarities was 6.6 (SD = 2.16), and the total score for Differences was 4.9 (SD = 2.2). Because each of the 10 items on both subcomponents represented a choice among four alternatives, random guessing by the children would be expected to produce means of 2.5 for each subcomponent. The Differences mean of

4.9 is only 2.4 above this chance level. It is indeed a difficult test.

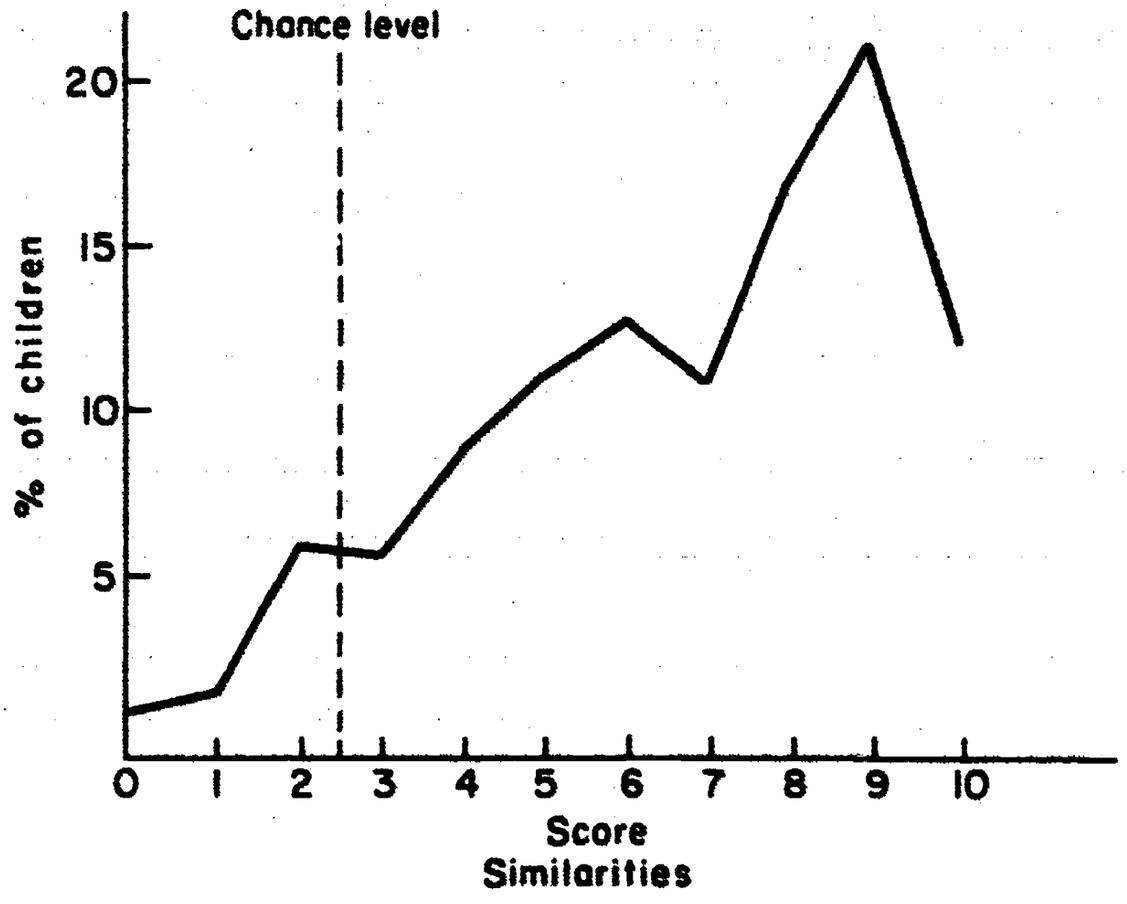
The distributions of the total scores for Similarities and Differences are presented in Figure 14. In it the Similarities total scores skew toward the higher scores, and the Differences total scores cluster more toward the middle and lower scores.

Item Analyses

The internal consistency of this component may be evaluated by inspecting the difficulty and discriminability of the items. Item difficulty is indicated by the percentage of children passing the item. It should be noted that random responding or mere guessing would result in 25% of the children passing the item by chance. Item discriminability is reflected in the correlation between success on the item and total score and in the average total scores of children who pass the item and those who fail the item.

Table 30 presents the findings related to difficulty and discriminability. Neither the Similarities nor Differences subcomponent possesses a desirable spread of difficulty, which would have a majority of items in the middle range and fewer difficult and easy items. The Similarities subcomponent is the better of the two. It had three items at the 60% level and three at the 70% level. The Differences subcomponent had two items with difficulty levels at chance expectancy, while the remaining items were skewed in the difficult direction. The range of difficulty of the items should be improved.

Regarding the discriminability of the items, all item correlations with the total score were positive. The majority of Similarities items



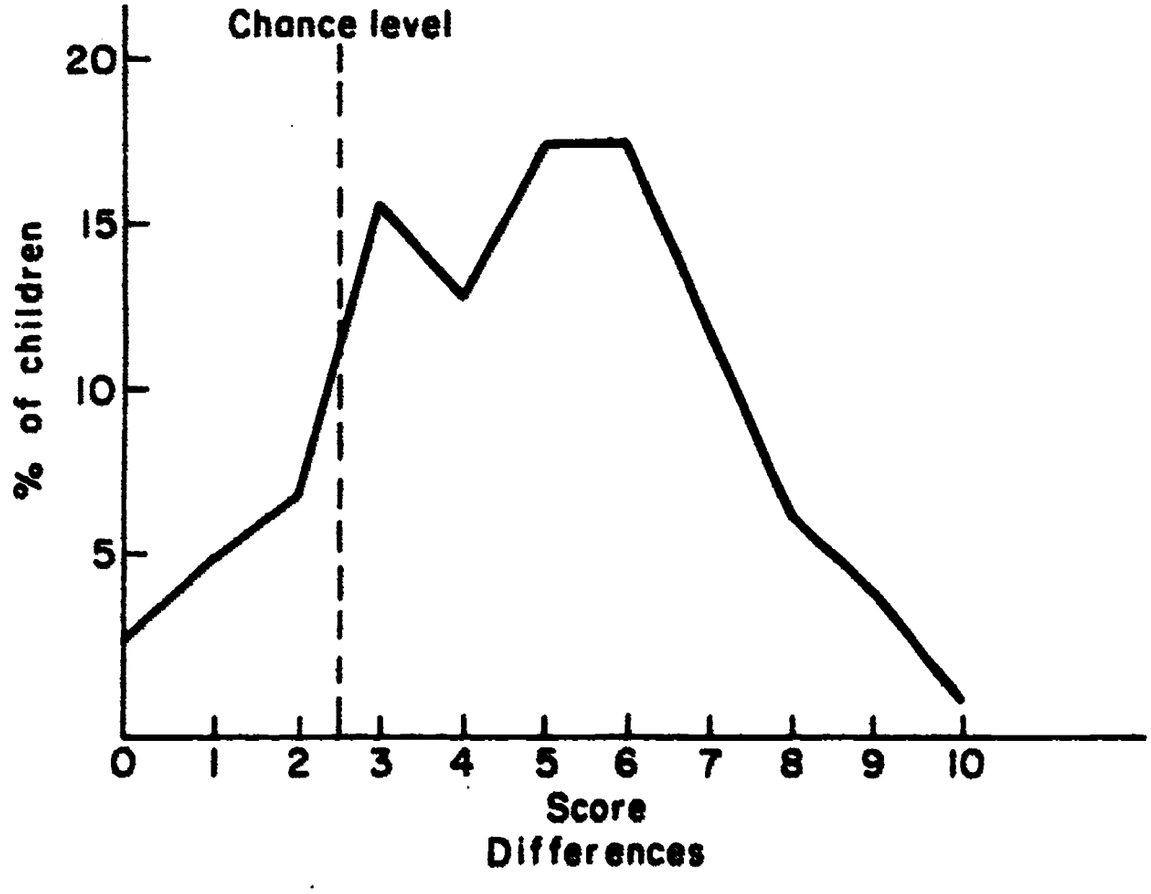


Figure 14. Distribution of visual discrimination scores.

TABLE 30

ITEM ANALYSES OF VISUAL DISCRIMINATION

Item (in order of decreasing difficulty)	% of children passing item	Biserial correlation with total score	Average total score of children who pass item	Average total score of children who fail item
		Similarities		
6	34	.56	8.2	5.8
9	44	.80	8.4	5.2
1	60	.77	7.9	4.8
10	65	.70	7.7	4.7
7	66	.87	7.9	4.2
5	70	.86	7.8	4.1
4	72	.88	7.7	3.9
3	79	.76	7.4	3.9
2	85	.66	7.1	3.9
8	85	.66	7.1	3.6
		Differences		
9	25	.45	6.1	4.5
8	27	.30	5.7	4.6
6	40	.45	5.8	4.3
3	43	.73	6.4	3.8
10	43	.72	6.4	3.8
5	46	.70	6.2	3.8
7	53	.58	5.9	3.8
4	63	.64	5.7	3.5
2	69	.82	5.8	2.9
1	83	.68	5.4	2.5

possess a good relationship to the total score. Eight out of 10 correlations were in the .70 and .80 range, and the average of all 10 was .76. The correlations of Differences items with the total score were more modest. They averaged .61. Further indication of the discriminability of each item is given in the figures which allow comparison of the total scores of children who passed the the item and of children who failed the item. These differences were all statistically significant beyond the .01 level.

As assessed by the Kuder-Richardson formula 20, the interitem consistency of Similarities was .78 and that for Differences was .61. Both subcomponents are thus indicated to be relatively homogeneous Similarities more so than Differences.

Subgroup Analyses of Total Scores

Table 31 presents the findings regarding the age, sex, and race subgroups. Age-related differences for Similarities and Differences show scores that significantly differ (for Similarities, $F = 25.17$, $df = 2/264$; for Differences, $F = 14.74$, $df = 2/245$). Tukey's tests, following these analyses of variance, indicated that a significant difference existed between each age subgroup mean ($3 < 4 < 5$ in each case). Thus, as expected for a cognitive/perceptual skill, Visual Discrimination scores increased with age. Confidence in the measure is supported in this regard.

No sex or race differences were expected, but one occurred, that for race on Similarities ($t = 2.17$, $df = 149$). This result seems uninterpretable at this time, particularly as a parallel result did not obtain for Differences.

TABLE 31

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR TOTAL SCORES OF THE VISUAL DISCRIMINATION COMPONENT

Group		N or n	$\bar{X}$	SD	p
Similarities					
All children		267	6.6	2.6	
Age:	3	59	5.0	2.5	
	4	136	6.6	2.4	<.001
	5	72	8.0	2.1	
Sex:	Male	136	6.7	2.7	ns
	Female	131	6.6	2.4	
Race:	White	71	7.5	2.3	<.05
	Black	80	6.6	2.6	
Differences					
All Children		248	4.9	2.2	
Age:	3	47	3.7	2.1	
	4	130	4.8	1.9	<.001
	5	71	5.8	2.4	
Sex:	Male	125	4.9	2.4	ns
	Female	123	4.9	2.1	
Race:	White	70	5.4	2.2	ns
	Black	76	4.8	2.4	

Recommendations

The following recommendations are made regarding future use of the Visual Discrimination component:

a. For any further work with this measure, the drawings should be corrected so that stimuli intended to be similar are completely identical. Minor unintended differences must be removed.

b. Randomization of the positions of correct responses must be achieved.

c. Similarities: The psychometric properties of this subcomponent, while not excellent, are probably adequate for its continued inclusion in the Inventory. It would be desirable, however, to achieve a better balance and spread of difficulty levels among the items.

d. Differences: This subcomponent requires further development before its further use can be recommended. Instructions for the children must be more successful in conveying the task to be performed. Additional trial items are suggested as a means of teaching the task. New items need to be developed which will improve the difficulty range and discriminability of the items.

CHAPTER 10

SOCIAL ROLES

The Social Roles component was directed toward assessing how the child perceives aspects of (a) his role at a future time, i.e., when he grows up, and (b) the roles of significant others (mother, father and teacher) in his present environment.

There are four questions to which the child's verbatim responses are recorded. The first question, tapping children's perceived future roles and vocational aspirations, is a variant of the familiar, "What do you want to be when you grow up?" question:

> A (boy, girl) can be all sorts of things when (s)he grows up. What would you like to be when you grow up?

Contingent upon a child's responding, reinforcement is given and additional responses are sought:

> That's very interesting. Maybe you've thought of some other things you could be when you grow up. What else could you be when you grow up?

Following any additional responses, the child is asked, regarding his initial choice,

> Why do you want to be a(n) _____?

The remaining three questions asked for the child's perception of three social roles:

> What does a mother do? (What does a mommy do?)
>
> What does a father do? (What does a daddy do?)
>
> What does a teacher do?

Administration and Scoring

The examiners' feedback indicated that, while they thought these were "good" questions, the placement of them in the Inventory component could be improved. It was recommended that this component, the second to be presented, should be held until later. They felt that, for shy children in particular, there was a difficult transition from the nonverbal Dog and Bone component, the first component of the Inventory, to the heavy verbal requirements of the Social Roles component. It was suggested that the nonverbal-to-verbal transition be made more gradual by utilizing as the second component another nonverbal component such as Body Parts or a component such as Picture Naming, which requires less complex verbal responses.

The children's responses were scored for number of responses given (vocational aspirations question only) and for category of response context. Response categories were developed from pilot results, a priori judgments, and data from field assessment. Interscorer reliability as assessed by two independent scorings of 20 records was over 99% for number of responses and averaged 95% for category of responses.

Findings

The findings for the vocational aspirations question and for the three social roles questions are presented separately. Each of these areas yielded results which strongly suggest the value of gathering further data so that the impact of group care on these perceptions of children and their developing concepts of others' roles and their own future roles can be understood.

The Vocational Aspirations Question[4]

The children's responses to the vocational aspirations question were categorized into the following six categories:

1. All adult: a category composed of the following three subcategories which define its scope:
   a) Specific Occupation.
   b) Adult: nonoccupational adult status, e.g., "be a man."
   c) Parent.

2. Older child: attributes of older, bigger children, e.g., "be a girl scout," "be taller."

3. Same child: responses indicative of lack of projection into more mature roles, e.g., "Be a boy just like I am."

4. Fantasy: fictitious characters or roles no longer existing to any significant degree in contemporary society, e.g., Batman, princess, cowboy.

5. Nonhuman: animals and inanimate objects, e.g., doggie, tiger, and "a bath tub, so I could drink lots of water."

6. Other: responses not classifiable into the preceeding categories and not sufficiently frequent to warrant additional categories.

Eighty-eight percent of the children gave scorable replies (those other than don't know, no reply, etc.) to the question. The ability to

---

[4]These results were presented at the March, 1973, biennial meeting of the Society for Research in Child Development, Philadelphia, Pa. The paper, "What do you want to be when you grow up? Vocational Choice in Children Aged Three to Six," co-authored by E. P. Kirchner and S. I. Vondracek, forms the basis of the presentation in this section and is available separately upon request from the Pennsylvania Day Care Study Project, Amy Gardner House, The Pennsylvania State University, University Park, Pa. 16802.

give scorable responses was not related to the child's age, sex, or race. A mean of 2.0 responses were given (SD = .83). The number of responses was not related to age, sex, or race.

Table 32 shows the distribution of responses among the six response categories for the total sample. The largest percentage of children gave responses categorized as all adult; within that category, the mention of a specific occupation was most frequent.

TABLE 32

OCCUPATIONAL ASPIRATION RESPONSES OF THE TOTAL SAMPLE

Response category	Percent of children expressing aspiration[a]
All adult	69.1
Specific occupation	57.0
Adult	11.2
Parent	17.9
Older child	17.7
Same child	10.0
Fantasy	10.0
Nonhuman	20.1
Other	12.0

[a] N = 249

Chi-square analyses were first performed to determine if overall dependences existed between subgroup divisions and all the main categories of response except "other," which was omitted from all subgroup analyses.

When these overall chi-square analyses proved significant, further chi-square analyses were performed to identify which individual response categories showed significant differences within the age, sex, and race subgroups.

Age-related findings are summarized in Table 33. The all adult category increased significantly with age and within the all adult category specific occupational aspirations increased markedly. There were decreases in the nonhuman, older child and same child responses.

TABLE 33

AGE COMPARISONS OF OCCUPATIONAL ASPIRATIONS

	Percent of children expressing aspiration			
Response category	Age 3 yr[a]	Age 4 yr[b]	Age 5 yr[c]	χ^2 probability
All adult	57.4	65.9	84.1	.004
Spec. occ.	42.6	51.6	78.3	.0001
Adult	11.1	13.5	7.2	.41
Parent	14.8	19.0	15.9	.74
Older child	24.1	19.8	8.7	.06
Same child	14.8	11.1	4.3	.14
Fantasy	13.0	8.7	10.1	.69
Nonhuman	33.3	20.6	8.7	.003

[a] n = 54

[b] n = 126

[c] n = 69

These findings suggest that a major aspect of vocational development in early childhood is projection into one's future role as an adult, specifically an adult identified occupationally. The child in this period is seen as coming to conceptualize himself as one day achieving adult status and to conceptualize adults as having a particular role characteristic, i.e., engagement in an occupation.

These findings support contentions that the young child's developing cognitive process enables him to think of himself in a propositional manner--to consider not simply who he is now but who he might become. They are also consistent with Erikson's (1963) view that early childhood is a period during which the child explores what type of person he might wish to be in the future.

The age-related findings, however, are most relevant to Havighurst's (1964) schema. His first stage of vocational development begins at year 5, the age of our oldest subjects, and extends through year 10. During this stage, called "Identification with a worker," the child's primary task is to identify with significant adults and to integrate the concept of work into his ego ideal. The results of this study suggest an important precursor of this stage: mastery of the task of projecting oneself into the future and conceiving of oneself as one day achieving adult status. Once this task has been mastered, the child may begin the process of identification with significant adult workers.

Sex comparisons are summarized in Table 34. It can be seen that boys and girls equally frequently gave responses classified as all adult; they did not appear to differ in the ability to project themselves into the vocational future. The pattern of this vocational

projection, however, differed for boys and girls. The boys were more likely to project in terms of an adult role in general, and the girls were more likely to project in terms of the specific role of parent.

TABLE 34

SEX COMPARISONS OF OCCUPATIONAL ASPIRATIONS

Response category	Percent of children expressing aspirations		χ^2 probability
	Males[a]	Females[b]	
All adult	72.5	65.3	.27
Spec. occ.	58.8	55.1	.65
Adult	15.3	6.8	.06
Parent	6.1	29.7	.0001
Older child	12.2	23.7	.03
Same child	9.2	11.0	.78
Fantasy	13.7	5.9	.07
Nonhuman	19.1	21.2	.80

[a] $n = 131$
[b] $n = 118$

Boys and girls mentioned specific occupations with equal frequency. It was of interest, however, to analyze the specific occupation data further, inasmuch as previous research with older children (Looft, 1971a; 1971b) suggests that girls learn early that certain adult statuses are not open to them and that boys generally perceive a wider range of available occupational alternatives. A foreclosure of occupational options has been postulated to occur earlier for females than for

males. Figures 15 and 16 show the specific occupations mentioned by boys (Figure 15) and by girls (Figure 16). These plots show vividly how the range of available occupations appeared wider to the boys; their choices were more evenly distributed, more "scattered." The plot for the girls' choices shows clustering toward two occupations, nurse and teacher, which are the most common choices of older girls. Comparison of the two plots offers support for the notion of earlier foreclosure for girls, "earlier" in this case being the preschool years.

It has also been suggested that girls are more reality bound than boys. Our findings of more frequent older child responses by girls and more frequent fantasy responses by boys support this notion. Girls are more likely to think in terms of being a girl scout or going to high school; boys are more likely to think in terms of being Batman or Santa Claus. Here again, whether nature or nurture is at the root of these differences, boys' range of perceived possibilities seems less restrained.

It is noted that Figures 15 and 16 show instances of children, particularly girls, naming occupations traditionally associated with the opposite sex (e.g., a girl choosing fireman, a boy choosing nurse). Further research might explore whether these instances represent a liberalization in the availability of occupations to both sexes or merely indicate that these children have not yet learned prevailing sex role occupational stereotypes.

Racial comparisons are summarized in Table 35. Black children tended to give fewer responses than white children in the all adult

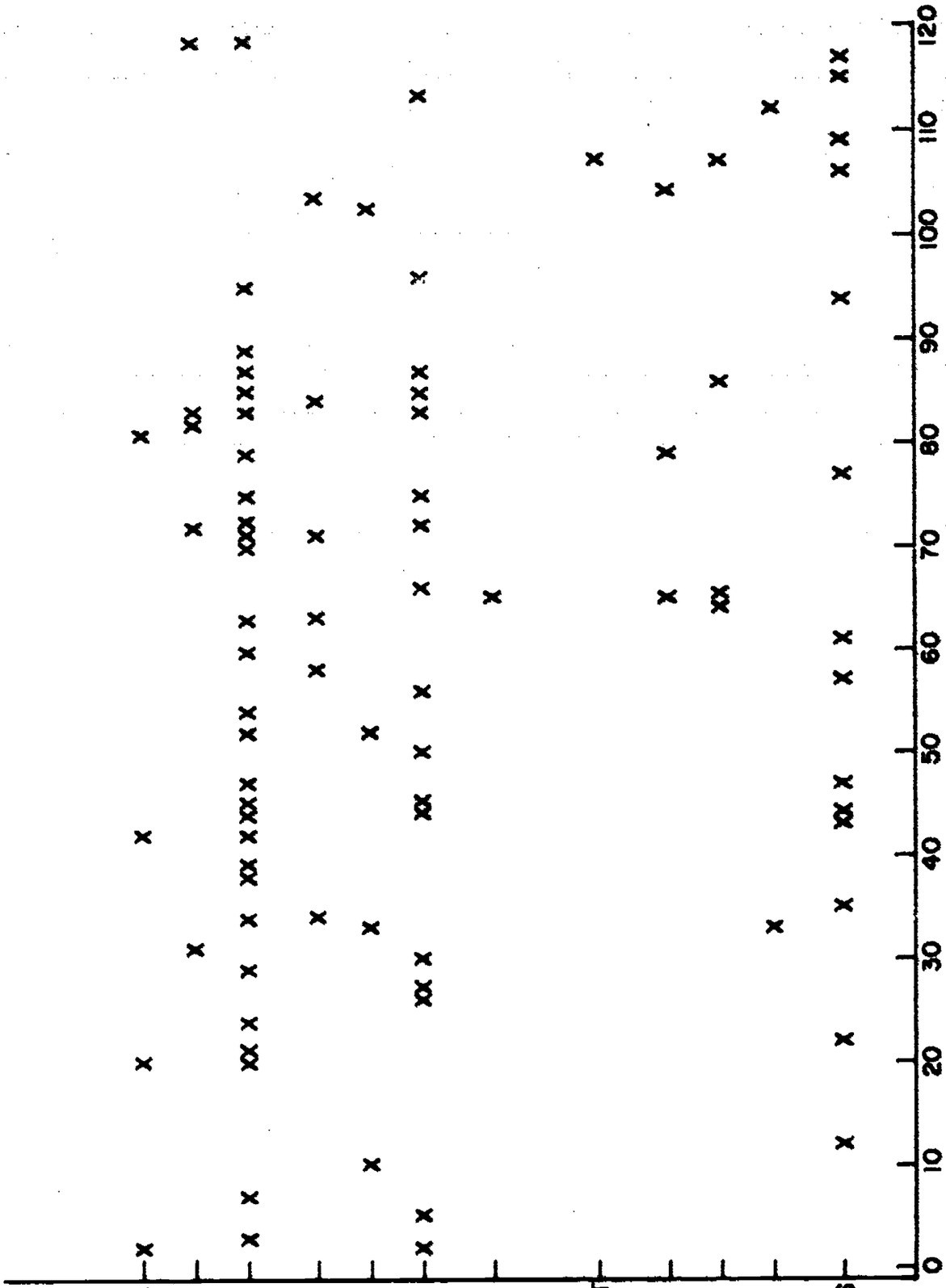


Figure 15. Occupational aspirations of boys.

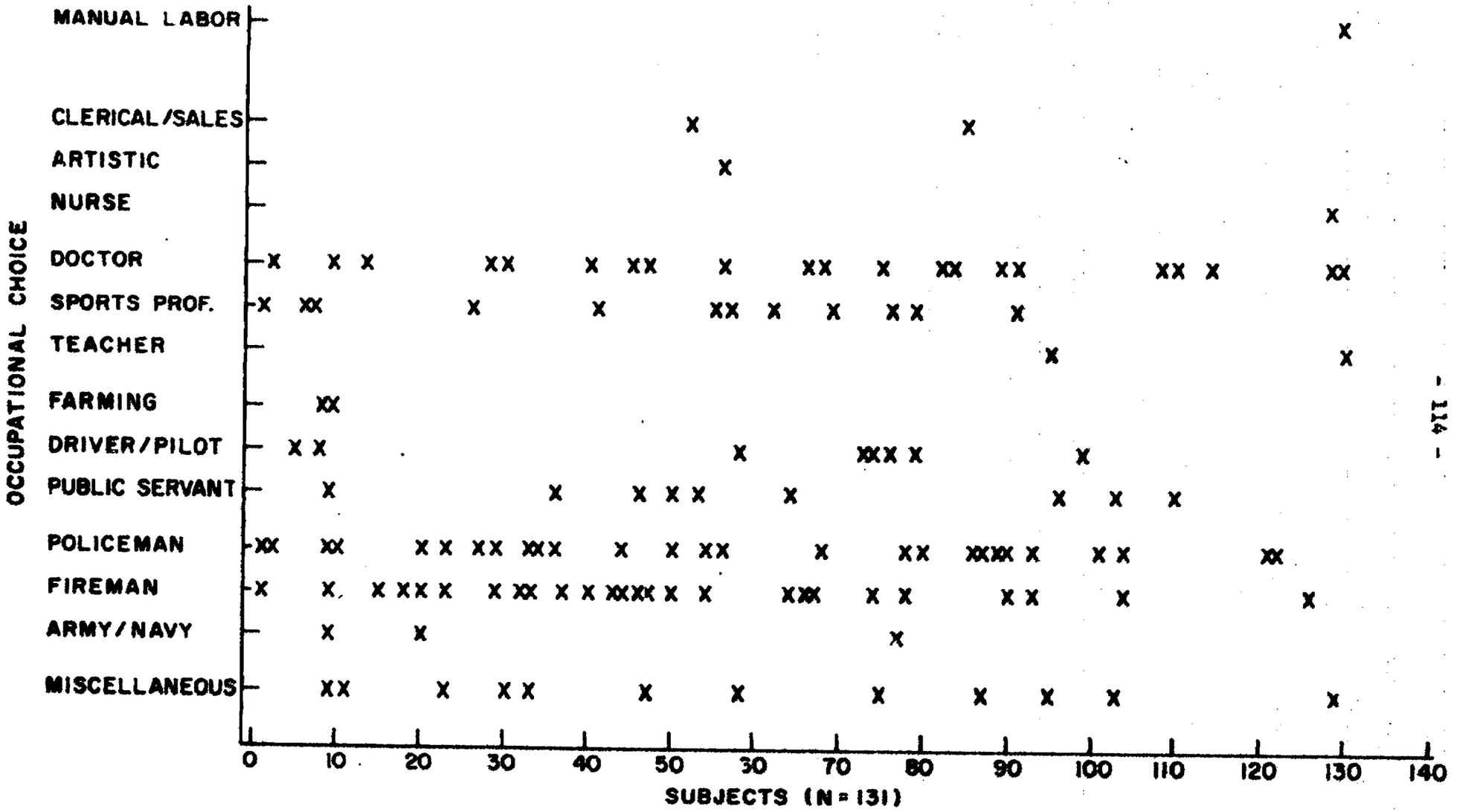


Figure 16. Occupational aspirations of girls.

TABLE 35

RACE COMPARISONS OF OCCUPATIONAL ASPIRATIONS

Response category	Percent of children expressing aspirations		χ^2 probability
	White[a]	Black[b]	
All adult	81.2	65.8	.06
Spec. occ.	71.0	41.9	.03
Adult	4.3	19.0	.01
Parent	20.3	15.2	.55
Older child	8.7	22.8	.04
Same child	4.3	16.5	.04
Fantasy	8.7	11.4	.78
Nonhuman	18.8	8.9	.13

[a] $n = 79$
[b] $n = 69$

category. Within this category, black children gave fewer specific occupation responses and more adult responses. Black children also gave more responses in the same child and older child categories. These findings suggest that black children are less mature in their vocational development from the standpoint of mastery of the task of projection into the vocational future. They are less likely to project themselves into adult statuscs, and their conceptualization of adult role less frequently involves occupational specificity. These racial differences are not related to paternal employment or presence in the home, as there were no race-related differences in these variables in

the sample. These findings are congruent with findings of racial disparity in the vocational planning of older children (Ansell & Hansen, 1971). They are also consistent with indications that black adolescents, in contrast to whites, see their occupational future as virtually predetermined and as involving limited, and primarily undesirable, possibilities (Hauser, 1971). Further, black children are more frequently admonished to avoid growing up to be like specified adults in their social sphere. It is suggested that the transmission of a sense of vocational predeterminism, together with an emphasis upon models and aspects of adulthood to be avoided, can serve to inhibit young black children's vocational projections as early as the preschool years.

In conclusion, the preschool period can be seen as a bona fide stage of vocational development; during these years the child comes to conceptualize himself as one day being an adult and having an occupation. Moreover, it is an important stage of vocational development during which population subgroup differences are already apparent. The sex and race differences found in these early years suggest that efforts toward achieving equality of vocational opportunity should not neglect the very young.

The Three Social Roles Questions

In order to compare the child's perception of three social roles, those of mother, father and teacher, as well as to describe them separately, the same categorization system was utilized for all three roles. The six categories were the following:

1. Works: responses which referred to a specific job such as "Drives a truck" and those which were nonspecific such as "Goes to work" or merely "Works."

2. Food: responses related to food preparation and serving.

3. Household: nonspecific references to household work (e.g., "Works around the house"), as well as specific masculine and feminine household chores such as "Washes dishes" and "Fixes the washer."

4. Personal: personal activities, recreations and pleasures such as grooming, resting, bowling, watching television, etc.

5. Child-Related: responses which identified some interaction with the child, such as teaches, directs, punishes, loves, helps, etc.

6. Other: responses which were unclassifiable in the above categories.

For the social role of teacher, the child related responses were further categorized into the following four subcategories: (a) teaches, (b) directs, (c) succors, and (d) other.

Of the 282 children asked the questions, 90.4% gave a scorable response to the questions about the roles of mother and father, and 84.0% were able to provide an answer to the question, "What does a teacher do?" There were no significant age, sex, or race differences in the ability to give scorable replies.

A summary of the responses given by the children to the three social roles questions is presented graphically in Figure 17 and is discussed more fully in the following paragraphs in which only significant findings or trends are presented.

The children's responses showed significantly more food-related responses for mother than they did for father ($\chi^2 = 87.62$, $df = 1$, $p < .001$) or teacher ($\chi^2 = 84.6$, $df = 1$, $p < .001$). Mother also received more household responses than did father ($\chi^2 = 7.78$, $df = 1$,

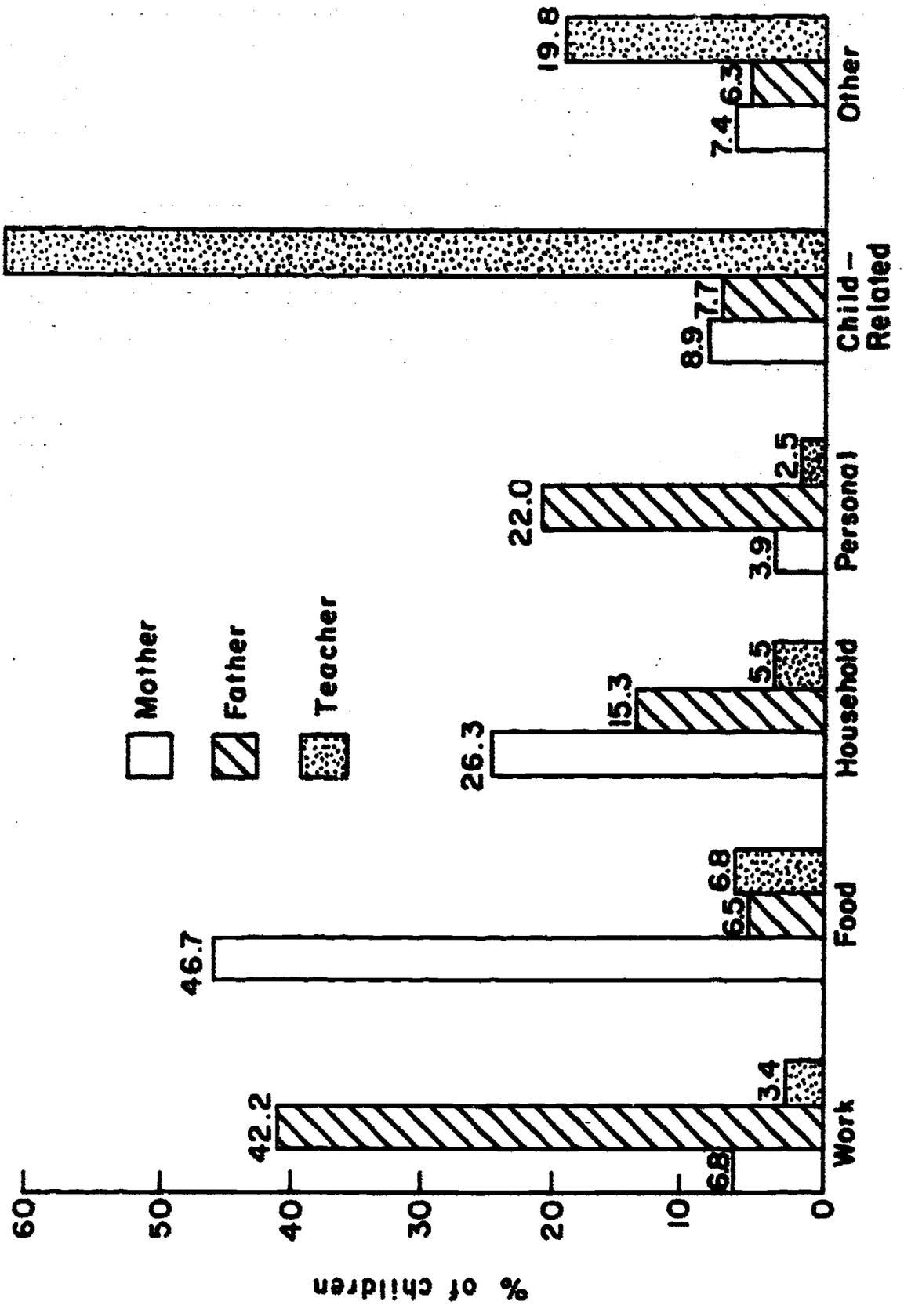


Figure 17. Categories of child responses for the role of mother, father and teacher.

$p < .01$) or teacher ($\chi^2 = 32.1$, $df = 1$, $p < .001$).

Father received more work-related responses than did mother ($\chi^2 = 81.047$, $df = 1$, $p < .001$) or teacher ($\chi^2 = 82.178$, $df = 1$, $p < .001$). Additionally, father received significantly more household responses than did teacher ($\chi^2 = 27.5$, $df = 1$, $p < .001$).

The role perception of teacher involved significantly more child-related responses than did that of mother ($\chi^2 = 109.8$, $df = 1$, $p < .001$) or that of father ($\chi^2 = 116.3$, $df = 1$, $p < .001$) and more responses classified as "other" than did the mother role ($\chi^2 = 16.2$, $df = 1$, $p < .001$) or the father role ($\chi^2 = 18.8$, $df = 1$, $p < .001$).

Thus responses to father were primarily in the work category (42.2% of the responses) and mother received primarily food (46.7%) and household (26.3%) responses. These findings are consistent with those of Finch (1955) and Hartley (1960), who found that father is seen primarily as the economic provider, and mother as a "homemaker." There are no previous data with which to compare the results of the role of teacher, but the results of the present study show that the teacher was overwhelmingly described in terms of child-related behaviors. Thus the role of father is described as one that involves primarily work and personal pleasures; the role of mother is seen as embodying food preparation and serving and doing household chores, and the role of teacher is described as one which involves much interaction with the child. The large number of responses which were classified as "other" for the teacher could be taken to indicate that this role is less well delineated by children.

Age-related findings regarding role perception of mother show that work responses increased with age ($\chi^2 = 10.2$, $df = 2$, $p < .01$), a

finding perhaps related to the child's enlarging conceptualization of work or his growing awareness of the mother's actual role behavior. Food responses fluctuated with age ($\chi^2 = 7.6$, $\underline{df} = 2$, $\underline{p} < .05$), being greater at age 4 than at ages 3 or 5.

In regard to age differences in the perception of the role of father, work responses also showed a significant increase ($\chi^2 = 8.6$, $\underline{df} = 2$, $\underline{p} = .05$). There were trends toward increasing personal responses ($\chi^2 = 4.1$, $\underline{df} = 2$, $\underline{p} = .10$) and toward decreasing food responses ($\chi^2 = 5.2$, $\underline{df} = 2$, $\underline{p} = .10$). Thus, with increasing age there seemed to be a developing consensus in the perception of father as primarily working, and when not working, engaging in some personal activities and pleasures.

With increasing age, the role of the teacher showed an increase in child-related responses ($\chi^2 = 9.8$, $\underline{df} = 2$, $\underline{p} < .01$), a decrease in food responses ($\chi^2 = 8.5$, $\underline{df} = 2$, $\underline{p} < .05$), and a decrease in "other" responses. Analyses of the subcategories of child-related responses showed that the increase with age in the child-related category was largely due to an increase in responses classified as teaches ($\chi^2 = 22.9$, $\underline{df} = 2$, $\underline{p} < .001$), there being no significant differences in succor, directs and other subcategories. Thus the perception of the role of the teacher also becomes more delineated with age, not surprisingly focusing upon the teaching aspects of the role.

There were a few interesting sex differences in the children's responses to two of the social role questions. While no significant sex differences were obtained for teacher, females tended to give more food-related responses for mother ($\chi^2 = 3.2$, $\underline{df} = 1$, $\underline{p} < .10$) and more personal responses for father, ($\chi^2 = 4.8$, $\underline{df} = 1$, $\underline{p} < .05$), and males gave more household responses for father ($\chi^2 = 6.7$, $\underline{df} = 1$, $\underline{p} < .01$).

Thus females tend to perceive the role of mother as involving more food preparation than do males. For the role of father, it appears that females view father as engaging in more personal activities, while the males see father as doing more work around the house.

It is intriguing to view these differential perceptions of mother and father roles in light of the cultural stereotype of the female complaining that the male does not do enough work around the house and the defense of the male that he does indeed do so. The data show that these differential views are apparent in the role perceptions of young children. Perhaps children's experiences with their same sex parent lead to such perceptions or perhaps the stereotypical attitudes influence their role perception; only further research can determine the origin of these perceptions.

There were no race-related findings regarding perception of the mother and teacher role. The only role perception which showed different responses by black children and white children was that of father. Here the black children gave more food responses ($\chi^2 = 4.40$, $\underline{df} = 1$, $\underline{p} < .05$), and the white children more household responses ($\chi^2 = 3.56$, $\underline{df} = 1$, $\underline{p} < .06$). Without further data it can only be speculated that these differential perceptions may represent differences in role behaviors of black and white fathers.

Additional analyses were performed to determine if the role perceptions were related to whether or not the mother was working outside the home and whether or not the father was present in the home. There were no significant differences in the three role perceptions between children from two-parent families and mother-only families. Likewise,

there were no significant differences in the perception of the mother and father roles which were related to whether or not the mother was working outside the home. The perception of the role of the teacher, however, seemed related in two instances to whether or not the mother worked outside the home. Children of employed mothers more frequently gave child-related responses ($\chi^2 = 6.62$, $df = 1$, $p = .01$) and less frequently gave "other" responses ($\chi^2 = 6.09$, $df = 1$, $p < .02$). The differences in child-related responses were largely due to differences in the teaches subcategory ($\chi^2 = 7.64$, $df = 1$, $p < .01$), there being no significant differences in any other subcategory of child-related responses.

Without further work, interpretation of these data is hazardous. Inasmuch as similar changes were observed to be related to age (and employment of mother did not increase with the age of the child), it may be that children of employed mothers have a more mature and delineated concept of the teacher's role. It could also be hypothesized that for children of employed mothers, the child-related aspects of the teacher's role become more salient.

In summary, these results give insight into the phenomenological world of the child. In the description of the perceptions that children have of the social roles of mother, father, and teacher trends in the development of these perceptions appear. Children as young as 3 years have begun to develop a perception of mother as one who prepares food and works around the house, of father as one who works and at other times engages in personal activities, and of teacher as someone who teaches and interacts with children. These perceptions become more

pronounced with increased age. Some, but not a great many, race and sex differences appear. As Dubin and Dubin (1965) have pointed out, the perceptions which children have of their environment should be considered when one attempts to explain their behavior. It is not simply the behavior of significant others to which children respond but their perceptions of others' behaviors.

Recommendations

This component involves questions which are easily administered to children although it should be preceeded by a sufficient "warm-up" to its highly verbal nature. The findings, already interesting in their own light, suggest the importance of investigating the impact of group care experiences and various kinds of group care experience on a child's vocational aspirations and on his perception of significant social roles. Further work should explore these areas.

CHAPTER 11

SOCIAL COMPETENCE QUESTIONS

The component of Social Competence Questions attempts to assess selected aspects of a child's developing social maturity. It utilizes the following five questions, the child's verbatim response being recorded by the psychometrist:

1. Let's pretend you just got up in the morning and want to get dressed. What do you do?

   Inquiry (a) Who puts on your shirt/dress?
   (b) Who buttons or zips your coat?
   (c) Who ties your shoes?

2. Let's pretend that you were playing and your toy broke. What do you do?

3. Let's imagine that you cut your finger. What do you do?

4. Imagine that a child much smaller than you are starts to fight with you. What do you do?

5. Why do people have to work?

Administration and Scoring

Examiners reported that the children were, for the most part, able to understand and answer these questions. It was felt, however, that the "wordiness" of two questions (get dressed and starts to fight) should be reduced.

The children's responses were tabulated according to content categories established to best represent the range of responses the children gave. Percent agreement between two independent scorers judging 30 records averaged 96% (range 90% to 100%) for the various tabulations made.

Findings

The Get Dressed Question

Of the 277 children asked the question concerning getting dressed, 251 (90.6%) were able to give a scorable reply. There were no sex or race differences in ability to respond. Age differences, however, tended toward significance ($\chi^2 = 4.35$, $df = 2$, $p = .11$); 84.4% of the 3-year-olds, 91.4% of the 4-year-olds, and 94.5% of the 5-year-olds gave scorable replies.

Table 36 presents the tabulation of the children's responses. The most frequent response (doing something other than getting dressed) indicates that the children were not able to focus upon the intent of the question, getting dressed. This finding was assumed to be a function of poor wording of the question. In regard to those children who did respond in terms of getting dressed, most children omitted spontaneous reference to any help in doing so. There were no significant sex or race differences in these categories. Age differences were significant ($\chi^2 = 22.90$, $df = 8$, $p = .004$) but showed no consistent increase or decrease with age in any of the categories. The do something else category, for example, had a frequency of 37.0% for 3-year-olds, 52.3% for 4-year-olds, and 40.6% for 5-year-olds.

TABLE 36

RESPONSES TO THE QUESTION,

"LET'S PRETEND YOU JUST GOT UP IN THE MORNING

AND WANT TO GET DRESSED. WHAT DO YOU DO?"

Response	Percent of children responding[a]
Do something else (e.g., eat breakfast, play)	45.8
(I) get dressed, unelaborated	35.9
(I) get dressed, elaborated (e.g., "I put on my shirt and socks.")	15.1
Get dressed, help from others specified	0.8
Grooming only (e.g., "I brush my teeth.")	2.4

[a] N = 251

Responses to the three-question inquiry regarding who puts on the dress or shirt, buttons or zips the coat, and ties the shoes are summarized in Table 37. The child himself and his mother were by far the most frequently specified persons. It must be stressed, of course, that the relationship of the children's responses to actual dressing behaviors is not known. The child's response that he himself accomplishes these tasks may represent reality, the child's exaggeration of reality (A child may be able to put on his shirt or may do so occasionally and reply to our question as though it is his usual behavior.), or the child's perception of the ideal (He may think he should do these things and for a variety of reasons may report that he does them.). Significant or

nearly significant overall age differences were found in responses to each of these questions ($\chi^2 = 11.60$, $df = 6$, $p = .07$ for shirt; $\chi^2 = 21.02$, $df = 6$, $p < .01$ for buttons; $\chi^2 = 23.12$, $df = 6$, $p < .001$ for shoes). Table 38 presents the subgroup findings for the two predominant responses, self and mother. It can be seen in Table 38 that responses specifying self-help increased with age for each of the specific behaviors questioned.

Sex and race comparisons yielded no significant findings except that boys indicated more independence in putting on shirts than girls did in putting on dresses. It is felt that this finding and the lack of other sex-related findings is most parsimoniously interpreted as reflecting the differential complexity of the two garments.

TABLE 37

RESPONSES TO THE SPECIFIC INQUIRY OF
"WHO PUTS ON YOUR SHIRT/DRESS?,"
"WHO BUTTONS OR ZIPS YOUR COAT?," AND
"WHO TIES YOUR SHOES?"

	Percent of children responding		
Response	Shirt/dress[a]	Buttons/zippers[b]	Shoes[c]
Self	43.5	49.8	34.1
Mother	48.7	42.0	49.5
Father	4.3	4.3	9.8
Other	3.5	3.9	6.6

[a] N = 230
[b] N = 231
[c] N = 214

TABLE 38

AGE-RELATED CHANGES IN RESPONSES TO

THE SPECIFIC INQUIRY ON DRESSING

Response	Percent of children responding								
	Shirt/dress[a]			Buttons/zippers[b]			Shoes[c]		
	Age 3	Age 4	Age 5	Age 3	Age 4	Age 5	Age 3	Age 4	Age 5
Self	26.8	46.4	53.1	29.8	48.2	70.3	12.2	32.4	55.0
Mother	60.7	47.3	40.6	57.9	42.7	26.6	67.3	48.6	36.7

[a] N = 230
[b] N = 231
[c] N = 214

Further analyses will investigate the relationship between the child's response to these questions and indices of his actual self-help behaviors. At this point, these data can be summarized as showing that, whether achievement or aspiration toward achievement is reflected in the child's response, the transition from the help of others to self-help is indeed salient in this period. Additionally, it is the mother who is or who is perceived to be the primary source of help in these activities.

The Broken Toy Question

A scorable reply to the question of the broken toy was given by 86.6% of the children. Ability to give a scorable response was not significantly related to sex or race. It was, however, significantly related to age ($\chi^2 = 7.67$, $df = 2$, $p = .02$); 81.3% of the 3-year-olds, 84.3% of the 4-year-olds, and 95.9% of the 5-year-olds responded.

Table 39 presents the children's responses to this question. At least two dimensions are of interest here: one is that of the assumption of personal responsibility (e.g., fixing it vs. getting someone else to fix it) and the other relates to maturity of coping (e.g., crying vs. finding another toy to play with). In regard to the assumption of responsibility, it can be seen from Table 39 that about equal percentages of children state that they would fix it or get someone else to fix it. Nearly as frequently, the child's response simply stated another person would be informed (the assumption probably being that the other person would take care of the problem). As for the maturity of coping, few children would cry, hide the broken toy, or throw it away. Most children responded in a way that indicated they would, with or without aid, repair or replace the toy or go on to a substitute activity.

There were no significant sex or race differences in these categories of response, but age differences were significant ($\chi^2 = 30.23$, $df = 18$, $p = .04$). Decreases were found in "other" responses and in the throw it away category. An increase occurred, expectedly, in the fix it category. Although one might, on the basis of increasing self-reliance with age, expect decreases in eliciting the aid of others or simply informing others, both of these categories were increasingly represented with age. These latter findings raise the possibility that group care enhances a child's reliance upon others. Comparison of these results with responses of children with no group care experience would be highly informative.

TABLE 39

RESPONSES TO THE QUESTION,

"LET'S PRETEND THAT YOU WERE PLAYING AND YOUR

TOY BROKE. WHAT DO YOU DO?"

Response	Percent of children responding[a]
Fix it	16.7
Get someone to fix it	15.9
Tell someone (inform only)	13.8
Replace with a new one	10.9
Throw it away	10.5
Substitute activity: play with another toy	8.8
Cry	5.4
Hide it	.4
Get another toy (unclear whether replace with identical toy or substitute)	4.6
Other	13.0

[a] N = 239

The Cut Finger Question

Scorable responses to the question of a cut finger were given by 89.2% of the children, with no significant differences related to sex or race. Ability to give a scorable response was significantly related to age (χ^2 = 22.73, df = 2, p = .0001) with 4- and 5-year-olds having a higher percentage of scorable replies (95.7% and 90.4%, respectively) than 3-year olds (73.4%).

Table 40 presents the categorized responses to this question regarding a cut finger. It can be seen that the responses of roughly half the children imply that they would help themselves in this situation, although a surprising 16% appeared to feel the help of a physician was called for. Only small percentages gave responses which might be considered inappropriate or incomplete coping strategies (e.g., bleed, cry, ignore). No significant sex or race differences were found in the distribution of responses, but again significant age differences appeared ($\chi^2 = 41.15$, $df = 16$, $p = .0005$).

TABLE 40

RESPONSES TO THE QUESTION,

"LET'S IMAGINE THAT YOU CUT YOUR FINGER.

WHAT DO YOU DO?"

Response	Percent of children responding[a]
Self help (e.g., "Put a Band-aid on it.")	45.5
Go to the doctor	15.9
Get someone to help	12.6
Bleed	6.9
Cry	6.9
Tell someone (inform only)	4.5
Ignore	1.6
Example given	1.6
Other	5.3

[a] N = 246

Inspection of the age-related differences indicates that there was no increase with age in self-help responses, but there was an increase in eliciting the aid of others. Age-related decreases occur in the ignore, example given, and "other" categories. As with the previous question, the increase with age in responses reflecting reliance upon others is puzzling, a comparison of these data with responses of children not in group care would throw light on the question of whether or not early group care promotes increasing orientation toward others as help-givers.

The Starts to Fight Question

A scorable reply to the question concerning starting a fight was given by 89.2% of the children. No statistically significant sex or race differences were found in this variable. Age differences were statistically significant ($\chi^2 = 10.91$, $df = 2$, $p = .004$); 79.7% of the 3-year-olds, 89.3% of the 4-year-olds, and 97.3% of the 5-year-olds were able to give scorable responses.

Table 41 presents the categorized responses to this question. Over half of the children reported that they would hit back, and about one-fifth reported they would inform an adult. The remainder of the responses were distributed among various other coping strategies, each with relatively low frequency.

Subgroup comparisons of the category of response indicated no significant race differences. Sex- and age-related differences, however, were found. The significant overall sex differences ($\chi^2 = 26.21$, $df = 10$, $p = .004$) were related to the differential use by boys and girls of the hit back and inform adult categories. Boys more frequently than

girls reported they would hit back (64.6% vs. 43.7%, respectively), whereas girls more frequently than boys reported they would inform an adult (31.9% vs. 12.6%, respectively). Age differences ($\chi^2 = 37.13$, $df = 20$, $p = .01$) indicated an increase in the inform adult category (13.7% at age 3, 24.0% at age 4, and 24.3% at age 5) and a decrease in the "other" category.

TABLE 41

RESPONSES TO THE QUESTION,

"IMAGINE THAT A CHILD MUCH SMALLER THAN YOU ARE

STARTS TO FIGHT WITH YOU. WHAT DO YOU DO?"

Response	Percent of children responding[a]
Hit him back	54.5
Inform adult	22.0
Be injured	4.1
Go away	4.1
Try to prevent him	3.3
Tell him to stop	2.0
Elicit adult help	2.0
Ignore him	0.8
Get mad	0.8
Cry	0.4
Other	6.1

[a] N = 246

There were, however, no appreciable age-related differences in the various low frequency categories or, unexpectedly, in the hit back category.

Attention should be called to the lack of an age-related decrease in this response, inasmuch as the hit back response represents a socially proscribed strategy in the case of a smaller child starting the fight (and is a response scored as a failure on the Wechsler Intelligence Scale for Children and the Wechsler Preschool Primary Scale of Intelligence). With increased socialization one would expect fewer hit back responses. Further work should explore whether, among other possibilities, changing social mores or factors related to group care are reflected in these findings.

The Work Question

The question regarding why people have to work seemed somewhat more difficult for the children than the preceeding four questions, to each of which about 90% of the children were able to give scorable replies. Approximately three-quarters of the children (76.4%) were able to give scorable responses to the work question. As with the other questions, there were no sex or race differences in ability to respond. There were also no age-related differences, whereas the other four questions yielded significant age-related differences in the ability to respond.

Table 42 presents the responses of the children to the work question. About one-fourth of the children expressed the reason for a person's working in terms of an unelaborated "needing" or "having" to work. This category, together with that of "example given," is felt to reflect a vague nonspecific conceptualization of reasons for having to work. The remaining categories of earning money, accomplishment, and satisfaction, however, are relatively specific, focused reasons. From the data in Table 42 it appears that earning money is perceived as the salient specific reason for having to work.

TABLE 42

RESPONSES TO THE QUESTION,

"WHY DO PEOPLE HAVE TO WORK?"

Response	Percent of children responding[a]
They have to, need to (unelaborated)	26.5
Earn money (unelaborated)	18.0
Earn money for necessities	7.1
Earn money for extras	1.9
Accomplish things--get things done	16.1
Example given	9.5
Want to, satisfaction and pleasure	5.7
Other	15.2

[a] N = 211

Subgroup comparisons yielded no significant sex differences in perceived reasons for work. There were significant findings, however, with regard to age ($\chi^2 = 44.12$, $df = 14$, $p = .0001$) and race ($\chi^2 = 23.58$, $df = 7$, $p = .0014$). Age differences showed a decline in the unelaborated "have to" category (32.6% at age 3, 29.9% at age 4, 15.5% at age 5) and in the giving of examples (26.1% at age 3, 3.7% at age 4, 6.9% at age 5). Increases with age occurred in the earn money category (10.9% at age 3, 16.8% at age 4, and 25.9% at age 5) and in the accomplish category (8.7% at age 3, 14.0% at age 4, 25.9% at age 5). These age differences are seen as reflecting an increasing focus and specificity in reasons

for having to work. Race differences indicated more black than white children responding in the unelaborated "have to" category (32.4% vs. 24.1%, respectively) and in the satisfaction category (11.8% vs. 0%, respectively).

Summary and Recommendations

These questions should be retained, with possible reduction in the verbosity of the get dressed and starts to fight questions. They yield information which can profitably be utilized to explore the impact of group care experiences on several aspects of a child's social development, particularly reliance upon self and others.

CHAPTER 12

WHO LIKES YOU?[5]

The Who Likes You? question was designed to assess perceived sources of esteem. It goes right to the heart of its intent. The children are simply asked:

> Tell me who likes you. Tell me as many people as you can who really like you.

Examiners were coached to convey by their tones, emphasis, and pace a very warm and positive connotation to the concept of liking involved in the question.

Administration

Older children and adults might be expected to be somewhat surprised by this question, to feel anxious, and to respond defensively. Psychometrists' feedback revealed that such was not the case with the great majority of the sample children, even with those (7) children who responded that nobody liked them. Typical comments of the psychometrists were: "They don't find a great deal of embarrassment; they just rattle off names." A few said, 'Nobody,' but didn't feel uncomfortable about it." "Maybe they aren't as involved with it as we are." A few instances were reported, however, in which the examiner had felt a child's embarrassment or sadness at replying "nobody" and had responded with a hearty, "Well, I like you!"

---

[5]Results from this component form the basis of a more technical article prepared for journal publication. Copies of this article, "Perceived Sources of Esteem in Early Childhood," are available from the Pennsylvania Day Care Study Project, Amy Gardner House, Pennsylvania State University, University Park, Pa. 16802

A more common difficulty reported by the psychometrists was ascertaining the identity, role, or relationship of a specific person whom the child had mentioned. Sometimes general questions and specific questions alike failed to provide the information needed to determine if the response referred to a person or animal, a real or imaginary being, a friend or sibling. The basis for this difficulty is a normal developmental stage of many preschool children. In this stage, the Piagetian egocentric stage, the child does not realize that there is any other viewpoint than his own and operates on the assumption that everyone possesses his own knowledge and viewpoint. A child in this stage cannot comprehend, for example, that some other person does not know who "Noonie" is, and his rapport with that person may suffer when, not knowing that Noonie is a neighbor's baby girl, the person asks such silly questions as "Is he a friend?...Your brother,...A pet...?" A special aspect of this problem was stated by one psychometrist: "They would get indignant if you didn't recognize names referring to other children in the day care center."

Scoring

Scoring Categories

Responses to this question were first categorized into (a) usable responses, those which named one or more perceived sources of liking, and (b) unusable responses, those which represented no response, "don't know," or irrelevant responses indicating the question was not understood. Usable responses were then further categorized. Categories for responses to the Who Likes You? question were developed to represent (a) obvious categories of relationship, e.g., mother, child

friend; (b) categories of particular interest vis à vis the child's day care experience, i.e., day care center staff and day care center children, and public servants; and (c) categories suggested by the data themselves, e.g., fantasy or TV characters, and the mention of specific persons as not liking the child.

While for more extensive work a few broad categories may be most useful and economical, the first efforts with this instrument concerned a broad range of categories. Thus the following classification was adopted:

Mother or mother surrogate

Father or father surrogate

Grandmother

Grandfather

Sibling(s)

Other (extended) family: cousin(s), uncle(s), aunt(s)

Child friend, specifically identified as day care center child

Child friend, not specified as to whether day care child friend or other

Fantasy or television characters

Public servants

Examiner

Day care staff

Other adults

Self

Pets

"Everybody"

Other human - not identifiable as child or adult

Other - unable to be categorized in other person or animal category

Nobody ("Nobody likes me.")

A specific person mentioned as not liking child ("Johnny doesn't like me." "My sister doesn't like me.")

Also tabulated for each child was the total number of perceived sources of esteem mentioned. The total of 57 was arbitrarily assigned to responses (such as "everybody") which indicated an indeterminate number. When means were determined on the total number of perceived sources of esteem, the "57" category was omitted from the tabulations.

Interscorer reliability, assessed by the percent of agreement between two independent scorers on 40 randomly selected response records, was 100% for response quality, 99% for category placement, and 98% for number of sources mentioned.

Findings

Ability to Respond

All of the 282 children in the sample were administered the Who Likes You? question. Of this total sample, 92% were able to give one or more scorable replies. This percentage increased from 88% at age 3 to 92% at age 4, to 97% at age 5, although the overall differences were not statistically significant ($\chi^2 = 4.47$, $df = 2$, $p = .11$). The remaining percentages in each case included the no response, don't know, or irrelevant response categories. Thus, this item is satisfactory in terms of appropriateness for this age group: the large majority of the children were able to handle this question. There were no significant sex or race differences in the percentage of children able to respond.

Number of Esteem Sources Mentioned

Table 43 presents summary data on the total number of sources mentioned. A mean of 4.3 sources was mentioned by the "average child" in the sample (SD 2.5). Girls mentioned significantly more sources than did boys (4.8 vs. 3.8).

TABLE 43

NUMBER OF SOURCES OF PERCEIVED ESTEEM MENTIONED

Group	N	Mean	p
All children	228	4.3	
Age: 3	51	3.8	
4	119	4.4	ns
5	58	4.6	
Sex: Male	118	3.8	<.01
Female	110	4.8	
Race: White	62	4.5	ns
Black	72	4.6	

Although the total number of sources increased with age, it was not a significant trend. A two-way analysis of variance of total number of sources indicated the significant effect of sex ($F = 10.96$, $df = 1/216$, $p = .001$), the nonsignificant effect of age ($F = 1.83$, $df = 5/216$, $p = .11$) and the nonsignificant age by sex interaction ($F = 2.02$, $df = 5/216$, $p = .08$). The possible age trend in the total number of sources mentioned was also investigated by correlating the child's age in months

with the total number of sources mentioned. The resulting low correlation of .13, while statistically significant at the .05 level (df = 227), indicated no appreciable relationship between age and number of sources mentioned.

The finding that the total number of sources mentioned was minimally, if at all, related to the age of the child eases the concern that the total might have been related to a child's verbal facility, memory span, or associational repertoire. If such were the case, the total would be expected to increase with age. It does not; therefore, there is some indirect evidence for divergent validity. Whatever is being measured does not exhibit characteristics expected of a verbal or memory skill item.

There were no significant race differences in the number of sources mentioned.

Relationship Between Number of Esteem Sources Mentioned and Size of the Child's Family

It might be hypothesized that the size of a child's family is related to the number of esteem sources he mentions. When size of family and total number of sources mentioned are correlated, the r of -.14, while statistically significant (df = 227, p = .05), is so small as to be insignificant for practical purposes. The correlation represents a frail inverse relationship, i.e., if anything, the smaller the family, the more sources mentioned by the child.

Categories of Esteem Sources Mentioned

Table 44 presents the sources of perceived esteem mentioned by the 260 children who gave scorable replies to the Who Likes You? question.

TABLE 44

ESTEEM SOURCES MENTIONED

Source	Number of children mentioning source	Percentage of children mentioning source[a]
Child friend	136	52
Day care child friend	32	12
Other child friend	119	46
Sibling(s)	96	49
Mother	120	46
Father	86	33
Other (extended) family	43	17
Grandmother	34	13
Day care staff	31	12
Grandfather	24	9
Pets	18	7
Other adults (nonfamily)	16	6
Examiner	10	<5
Self	9	<5
Nobody	7	<5
Source of dislike	7	<5
Everybody	6	<5
TV, fantasy	5	<5
Public servants	1	<5

[a]Percentages based on N of 260 children except in case of sibling as source where N is restricted to those 196 children who were known to have siblings. Percentage total exceeds 100 because of multiple responses.

It should be noted that in the case of categories which could include the mention of several separate persons, it was only the use of the category which was tallied. A child who mentioned three siblings, for example, was tallied once as utilizing the sibling category.

Many categories were infrequently used. These categories might ordinarily be combined into an "other" category but were retained for presentation in order to make the information regarding their relative infrequency available. Thus the domain of sources should be adequately covered by the initial categorization. However, a category of "other" was indeed utilized, although not for the traditional catch-all of low frequency responses. The "other" category contained those 52 sources for which specification of adult, child, etc., categories could not be made, most probably because of the difficulties inherent in the egocentric stage of the subjects (see previous discussion of this issue on Page 141).

Apparent from the findings presented in Table 44 is the relative salience of child friends, siblings, and parents as perceived sources of esteem. These results clamor for comparative data from children without day care or preschool experience. Knowing to what extent these family and peer sources are mentioned by children cared for at home could tell us a great deal about the possible impact of day care on the early socialization of young children. It may be that the importance of the peer group, traditionally recognized and focused upon in discussions of the elementary school and adolescent child, is heightened in the child in group care. Without comparative data, speculation can only be continued.

These comments regarding the need for a control group apply equally to other categories, for embedded in these categories are data relevant to some broad day care issues. The parental categories are the best examples here. Critics of early group care for children often hypothesize that, "It weakens the family." While this study may point to the relative salience of mother, the critic may point to assumed diminished absolute salience of mother ("Only 46%------"), assuming non day care children would mention mother more frequently. The need for comparative data is paramount. Within the current data, however, one notes that the percentage of children mentioning day care staff as sources of perceived esteem is a modest 12%. It appears that in this regard mother is not being usurped. From another point of view, that which underscores the possible contribution of day care experiences to the growing self esteem of the child, one may question why so few children mention significant adults in their day care world in answer to the question, "Who likes you?"

Age, Sex and Race Comparisons of Esteem Sources Mentioned

Those sources of perceived esteem which were of considerable frequency and/or interest were analyzed for possible age, sex, and race differences. The results are presented in Table 45. Apparent in Table 45 is that sex and race were the only variables relating to statistically significant differences. Girls more frequently than boys mentioned family members (mother, father and sibling) as sources of perceived esteem. There was no complementary difference in the mention of child friends, however: the two child friend categories did not show a sex difference in frequency of mention. A significant race difference appeared only in the

TABLE 45

MENTION OF PARENTS, SIBLINGS AND PEERS AS ESTEEM SOURCES

Group	Esteem source											
	Mother			Father			Sibling(s)			Child friend(s)		
	n	N	%	n	N	%	n	N	%	n	N	%
All children	120	260	46	86	260	33	96	196	49	119	260	46
Age: 3	33	58	57	21	58	36	22	36	61	25	58	43
4	57	130	44	41	130	32	49	105	47	56	130	57
5	30	72	41	24	72	33	25	55	45	31	72	57
	$\chi^2 = 3.55$, $df = 2$, $p = .17$			$\chi^2 = .39$, $df = 2$, $p = .82$			$\chi^2 = 2.16$, $df = 2$, $p = .27$			$\chi^2 = 3.47$, $df = 2$, $p = .18$		
Sex: Male	50	132	38	33	132	25	42	102	41	77	132	58
Female	70	128	55	53	128	41	54	94	57	63	128	49
	$\chi^2 = 6.73$, $df = 1$, $p = .01$			$\chi^2 = 7.18$, $df = 1$, $p = .01$			$\chi^2 = 4.55$, $df = 1$, $p = .03$			$\chi^2 = 1.82$, $df = 1$, $p = .18$		
Race: White	27	69	39	21	69	30	18	48	38	40	69	58
Black	43	85	51	33	85	39	37	59	63	42	85	49
	$\chi^2 = 1.58$, $df = 1$, $p = .21$			$\chi^2 = .84$, $df = 1$, $p = .36$			$\chi^2 = 5.76$, $df = 1$ $p = .02$			$\chi^2 = .80$, $df = 1$, $p = .37$		

Notes.--n = number of children mentioning person as source of perceived esteem; N = number of children in subgroup.

mention of siblings as a source of perceived esteem (black children mentioned siblings more frequently than white children). Age differences, not statistically significant, indicated a trend toward the diminishing mention of mother and siblings and toward an increase in the mention of child friends.

Because of particular interest in the impact of day care personnel and peers, separate analyses were made of the mention of day care staff and day care friends as sources of perceived esteem. The results are presented in Table 46. The results presented in Table 46 indicate that the mention of day care staff and peers was not significantly related to the child's age or sex. Although a day care child was mentioned significantly more frequently by the white children than by the black children, there was no racial difference in mention of day care staff.

Further Analysis of the Mention of Mother and Father

Presented here are data related to two questions: (a) If one parent is mentioned, how likely is it that the other parent is mentioned? and (b) Is the mention of both, neither, or either parent related to whether the child comes from a two-parent, mother-only, or father-only family?

Figure 18 shows the frequency with which children in each type of family mention neither parent, both parents, only the mother, or only the father. Analyses of these data show that the mention of parents tends to be a both or none affair, whether or not both parents are present in the home. In mother-only families and father-only families, as well as in two-parent families, the children mention either both parents or neither parent far more frequently than they mention either parent alone.

TABLE 46

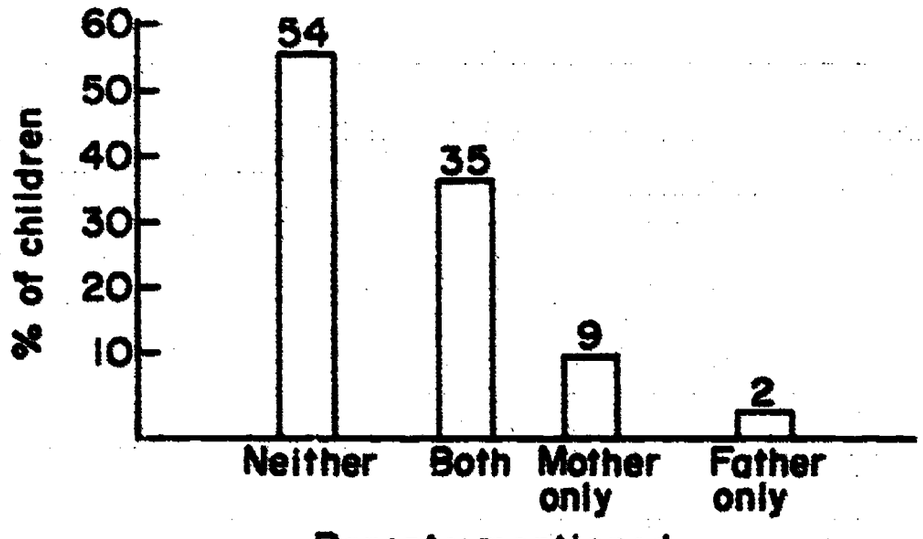
MENTION OF DAY CARE STAFF AND CHILD FRIENDS

AS ESTEEM SOURCES

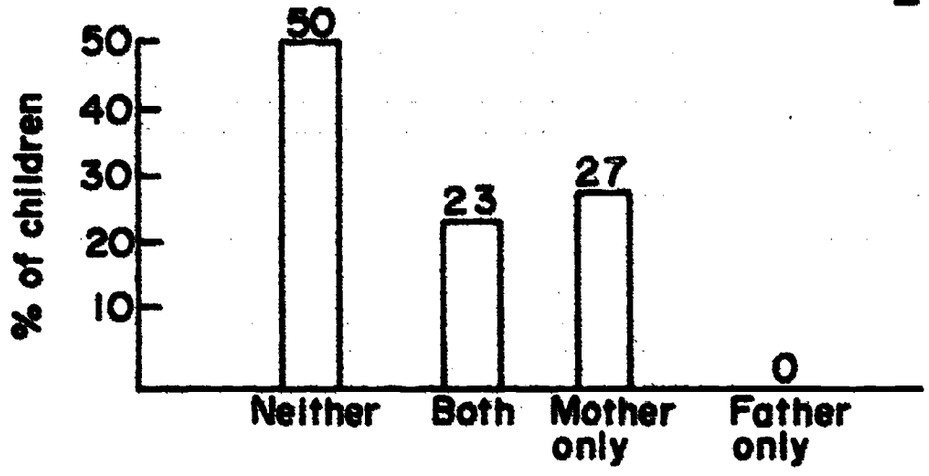
Group	Esteem source: Day care staff			Esteem source: Child friend from day care center		
	n	N	%	n	N	%
All children	31	260	12	32	260	12
Age: 3	9	58	16	3	58	5
4	16	130	12	17	130	13
5	6	72	8	12	72	16
	$\chi^2 = 1.61$, $df = 2$, $p = .45$			$\chi^2 = 4.07$, $df = 2$, $p = .13$		
Sex: Male	16	132	12	21	132	16
Female	15	128	12	11	128	9
	$\chi^2 = .01$, $df = 1$, $p = .93$			$\chi^2 = 2.58$, $df = 1$, $p = .11$		
Race: White	6	69	9	13	69	19
Black	12	85	14	6	85	7
	$\chi^2 = .62$, $df = 1$, $p = .43$			$\chi^2 = 3.86$, $df = 1$, $p = .05$		

Notes.--n = number of children mentioning person as source of perceived esteem; N = number of children in subgroup

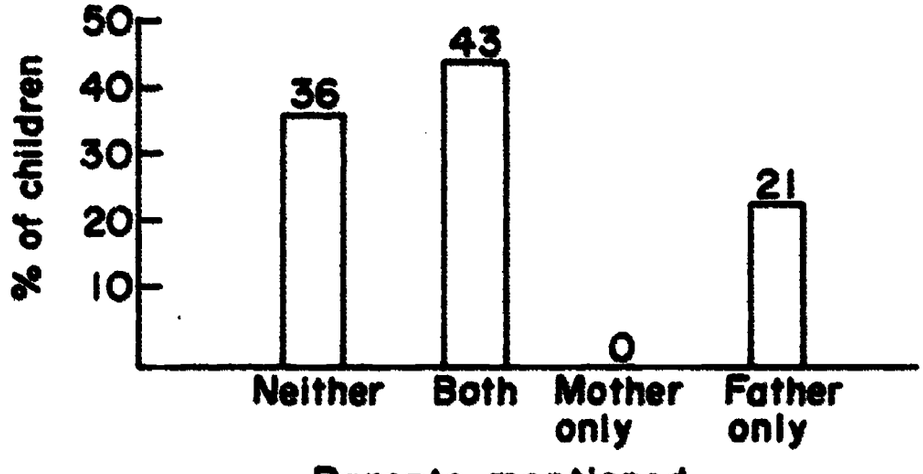
The significantly more frequent mention of neither or both parents than the mention of one parent only, regardless of family type, may be most parsimoniously explained as an instance of paired associate learning. The strength of the associative bond between the words mother and father probably over-rides differential percepts of parents as esteem sources in many young children in a test-like situation.



a. Two-parent families ($\underline{x}^2 = 87.93$, $\underline{df} = 1$, $\underline{p} < .001$)



b. Mother-only families ($\underline{x}^2 = 25.16$, $\underline{df} = 1$, $\underline{p} < .001$)



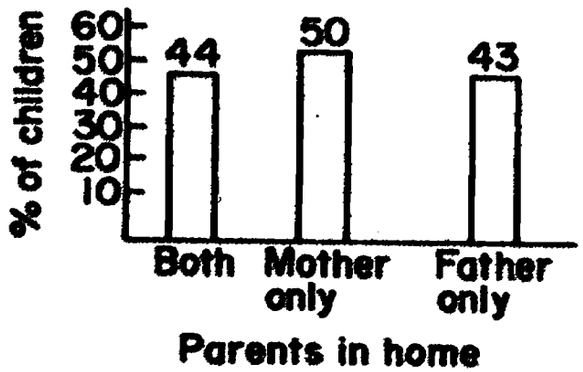
c. Father-only families (Fisher's Exact Test, $\underline{p} = .03$)

Figure 18. Children's mention of both, neither or either parent in answer to the question, "Who likes you?"

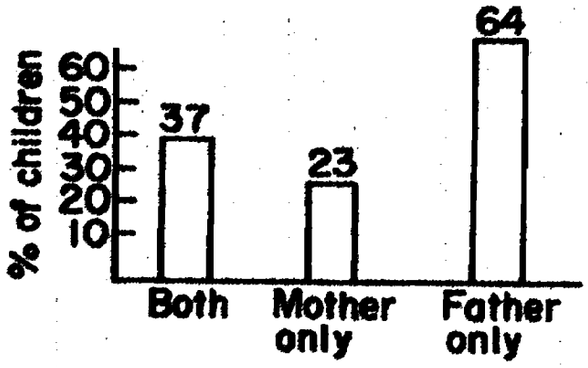
The magnitude of the percentage of children mentioning neither parent is a matter for empirical investigation, and the question of whether a group of nonday care children would give similar results should be studied. It has also been speculated that, had the question been phrased in terms of, "Who loves you?," parents might have been mentioned more frequently. This possibility is granted as a likelihood and deserves empirical attention, but the word "likes," spoken with the tones of "warm regard" that the psychometrists were trained to convey, comes closer to the variable that this study wished to measure. If asked, "Who loves you?," a child by reason of purely verbal associative learning could easily respond conventionally, "Mommy loves me," for example, and still not have conveyed in his response the sense of perceiving himself as really esteemed by his mother. It is felt that the present approach gets closer to perceived esteem and further from conventional, unreflective responses than having used the word "love" would have.

Results presented in Figure 19 relate the mention of mother, father, both, or neither parent to type of family (two-parent, mother-only, father-only). Here the mention of mother represents the total of mention of mother alone together with mention of mother in conjunction with mention of father. The mention of father is analogously defined.

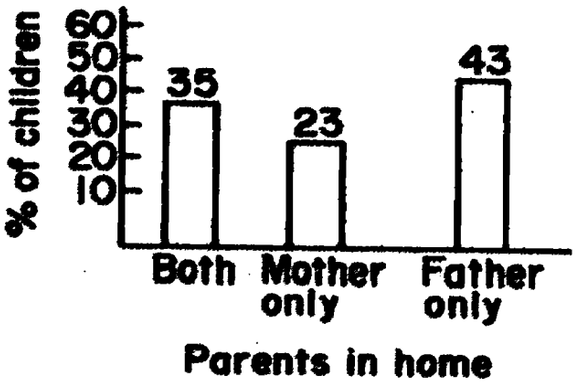
The only statistically significant relationship between parent(s) in the home and mention of parent(s) occurred in the case of the mention of father. He was mentioned most frequently in father-only families and least frequently in mother-only families. The mention of mother, however, was not significantly related to type of family. She was mentioned equally (within statistical error) in two-parent, mother-only



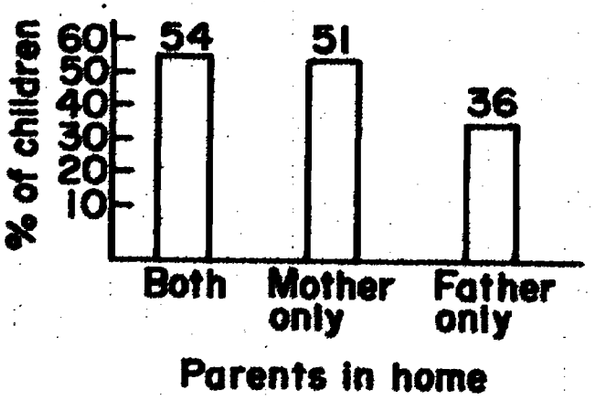
a. Mother mentioned ($\chi^2 = .66$, $df = 2$, $p = .72$)



b. Father mentioned ($\chi^2 = 11.83$, $df = 2$, $p = .003$)



c. Both parents mentioned ($\chi^2 = 5.19$, $df = 2$, $p = .08$)



d. Neither parent mentioned ($\chi^2 = 1.71$, $df = 2$, $p = .42$)

Figure 19. Mention of mother, father, both or neither parent by children from two- and one-parent families.

and father-only families. The mention of the absent parents in one-parent families (mother was mentioned in 43% of father-only families; father was mentioned in 23% of mother-only families) is a highly interesting finding. Perhaps mother is only a fond memory and father may visit once a month. To be a source of perceived esteem for a pre-school child may be a matter of the proverbial quality, not quantity.

Summary and Recommendations

The Who Likes You? component is easily administered and reliably scored. It yields highly interesting data relating to a child's perception of others in his world who can contribute to his developing self-regard. It should be retained in the Inventory, and further work should explore the impact of group care experience and different sorts of group care experiences on a child's number and type of perceived esteem sources.

CHAPTER 13

GOOD PERSON AND BAD PERSON

The Good Person and Bad Person questions were included in the Inventory in an attempt to understand certain aspects of a child's developing self-concept. It was reasoned that only the roots of what is called self-concept are being formed in children in the age range of this study. It was also reasoned that conceptualization of an "ideal self" preceeds actual self concept, that first a child develops the framework of what is desirable and not desirable, and only later construes his view of himself vis-à-vis this ideal. Thus, to assess the major evaluative dimension of the developing concept of ideal self the Good Person and Bad Person questions were used. The texts of these questions, presented at separate points in the administration of the Inventory, are as follows:

> Now I have another question for you. Let's think about what a good person does. We all try to be good people, don't we. What does a good person do?
>
> Now here's another question for you. We all try not to be bad people, don't we. (Pause) What does a bad child do?

A child's response is followed by a "prompt" asking for additional responses.

It should be noted that the questions are not strictly parallel in wording. As discussed in Volume I, the phraseology of "good person" and "bad child" resulted from pilot work indicating good comprehension and range of response to the present form of the questions.

Administration and Scoring

The psychometrists' reports regarding these questions were mainly general comments to the effect that the Good Person question was among the most difficult of items for the children but that the Bad Person question was much easier for them to answer. They felt, however, that rapport was not sacrificed for these questions and that a child could feel comfortable in saying, "I don't know."

Response categories were developed on the basis of responses obtained in pilot testing and field evaluation. The same categories were utilized for both the Good Person and the Bad Person question so that their relative use in the two questions could be assessed. A given category specifies only the content area of the child's response and not the direction or valence of the response, which is related to the particular question asked. An illustration may make this point clear. The responses, "Gets mad at his sister" (to the Bad Person Question) and "Doesn't get mad at his sister" (to the Good Person Question) are scored in the same content category, anger/hostility.

The categories utilized, together with some illustrative responses, are the following:

a. Minimal/vague. "Acts real good," "Does bad stuff" (unelaborated)

b. Plays. "Plays with toys"

c. Prosocial. "Helps people," "Doesn't share things"

d. Aggressive: a combined category composed of five specific categories:

1) Anger/hostility. "Doesn't stick out his tongue at people," "Gets mad and slams the door"

2) Aggressive. fights, hurts, "Doesn't hurt people," "Gets into fights"
3) Aggressive. kills
4) Aggressive to animals.
5) Destructive to objects.

e. Consequences: A combined category composed of three specific categories:

1) Consequences, corporal. "Doesn't get a spanking"
2) Consequences, tangible. "Gets candy"
3) Consequences, other. "Has to sit in a corner"

f. Rule conformity: a combined category of six specific categories related to rules, instructions, and prescriptions of social conduct:

1) Obedience and politeness. "Says 'thank you' and 'please'," "Doesn't do what his Dad says"
2) Eating-related behavior. "Eats all his food," "Spills his milk"
3) Cleanliness/neatness. "Cleans his room," "Gets all muddy"
4) Works.
5) Accomplishes. Reads, makes good paintings, dresses self
6) Lies/steals.

g. Other: A combined category of three categories:

1) Crying
2) Physical prowess and attractiveness. "Is strong," "Is real ugly"
3) Other: responses not classifiable in above categories.

Interscorer reliability, determined on the basis of two scorers independently categorizing responses from 40 children, was greater than 90% for all determinations made.

Findings

The results reported here are based upon the 279 children asked the Good Person question and the 275 children asked the Bad Person question.

Ability to Respond

Eighty percent of the children were able to give scorable responses to the Good Person question, and 87% were able to give scorable replies

to the Bad Person question. Both questions showed significant age effects in ability to answer the question (Good Person: $\chi^2 = 6.57$, $df = 2$, $p = .04$; Bad Person: $\chi^2 = 8.62$, $df = 2$, $p = .01$). On the Good Person question, the percentage of children giving scorable replies increased with age (71% at age 3, 77% at age 4, 89% at age 5), while for the Bad Person question fewer 3-year-olds (77%) gave scorable responses than 4-year-olds and 5-year-olds, (91% and 90%, respectively).

Neither question showed sex or race differences in the ability to give a scorable reply.

Number of Responses

The total number of responses given to the Good Person and Bad Person questions were about equal. The mean number of responses to Good Person was 2.4 (SD = 1.4) and to Bad Person was 2.3 (SD = 1.3). For each question, there was a significant relationship between age and number of responses given. For Good Person, 3-year-olds gave an average of 1.9 responses, 4-year-olds gave 2.2, and 5-year-olds gave 3.0 ($F = 10.80$, $df = 2/219$, $p = <.005$). For Bad Person, age-related results were strikingly similar to those for Good Person. Three-year-olds gave an average of 1.9 responses, 4-year-olds gave 2.1, and 5-year-olds gave 2.9 ($F = 11.07$, $df = 2/237$, $p = <.005$). Tukey's tests showed that for both questions the total number of responses given by 5-year-olds was significantly greater than that for 4-year-olds or 3-year-olds, the latter two groups not being significantly different in total number of responses given.

Mode of Response

Of interest was the children's mode of response, i.e., whether responses to each question were stated positively (e.g., does something), negatively (e.g., does not do something), or in the form of an example (e.g., "I was playing and some kid hit me."). It seemed possible that Good and Bad Person responses might differ in the relative frequencies of these modes of response and that subgroup differences might also be found.

Results indicated that Good Person and Bad Person responses were overwhelmingly stated in the positive mode (91% for Good Person, 94% for Bad Person). The negative mode was utilized infrequently (4% for Good Person, 5% for Bad Person) as was the use of examples (5% for Good Person, 1% for Bad Person). There were no differences in mode of response between Good Person and Bad Person responses, nor were there age, sex, or race subgroup differences in the mode of response for either question.

The high frequency of responses in the positive mode is impressive when the seeming relative frequency with which children are exposed to concepts stated in the negative mode (e.g., "A good girl doesn't stick out her tongue") is considered. One could also note that a major definition of a good person for Judeo-Christian culture, the Ten Commandments, utilizes 20% positive mode vs. 80% negative mode in the King James Version of the Bible.

Mention of a Specific Person in Response

In responding to the questions, some children spontaneously mentioned specific individuals as examples or, more frequently, as participants in the behavior they were describing. Although overall, these

responses were not common, they occurred with sufficient frequency to warrant frequency analysis. The mention of one or more specific persons occurred in 13% of the responses to the Good Person question and in 8% of the responses to the Bad Person question. (McNemar's test of this difference indicated a trend toward significance, $p = .07$.) These data should be interpreted only in a very tenuous fashion. The fact that descriptions of a good person more readily call a model of that behavior to a child's mind than do descriptions of a bad person might suggest that the Good Person concept is somewhat more concretely based during this developmental period.

Categories of Responses

Table 47 classifies the responses to the two questions into the specific and combined categories. The presentation of findings regarding the smaller specific categories is made here for the interested reader, but all further analyses utilized the combined categories.

When each question is considered separately, Table 47 reveals that the Good Person responses occurred most frequently in the rule conformity category (35%). The prosocial and "other" (16% each), the minimal/vague and plays (14% each), and the consequences and aggressive (less than 5% each) followed. Bad Person responses occurred predominantly in the aggressive (43%) category. This category was followed by consequences (23%), rule conformity (17%), "other" (9%), and minimal/vague, prosocial and plays (less than 5% each). Thus, Good Person is seen as defined most frequently by obedience to authority and to general standards of socially facilitating and gracious behavior. Bad Person is defined with more consensus and more specificity; most fre-

TABLE 47

RESPONSES TO GOOD PERSON AND BAD PERSON QUESTIONS

Category of response[a]	Question			
	Good person[b]		Bad person[c]	
	n	%	n	%
Minimal/vague	32	14	14	6
Plays	32	14	0	0
Prosocial	36	16	4	2
Aggressive:	2	1	104	43
Anger/hostility	0	0	4	2
Aggressive: fight/hurt	2	1	79	33
Aggressive: kill	0	0	4	2
Aggressive to animals	0	0	2	1
Destructive to things	0	0	15	6
Consequences:	9	4	58	24
Consequences: corporal	2	1	20	8
Consequences: tangible	4	2	6	3
Consequences: other	3	1	32	13
Rule conformity:	76	35	40	17
Obedience & politeness	19	9	31	13
Eating-related behavior	20	9	2	1
Cleanliness/neatness	18	8	4	2
Works	8	4	0	0
Accomplishes	10	4	0	0
Lie/steal	1	1	3	1
Other:	35	16	21	9
Crying	0	0	8	4
Physical prowess	2	1	2	1
Other	33	15	11	5

[a]Direction of child's response differs for Good Person and Bad Person tallies. For example, in the prosocial category a child's response to the Good Person question might be, "He's friendly," and to the Bad Person question might be, "He's not friendly."

[b]Percentages based on N of 222.

[c]Percentages based on N of 240; total of percentages exceeds 100 because of rounding errors.

quently a bad person is described as one who engages in aggressive behaviors.

The frequencies in each response category for Good Person and Bad Person were compared utilizing McNemar's test for the significance of changes. Results indicated significantly different frequencies in the case of each category. Probabilities ranged from .02 to <.0001.

The consequences category requires comment. This category was not expected on the basis of initial pilot work but appeared, as is apparent in Table 47, with considerable frequency during the actual assessment. It will be recalled that the questions asked, "What does a (good person, bad child) do?" The wording reflected our intent to assess the child's description of a hypothetical person's behavior, not the consequences of that behavior. Yet many children gave responses, such as "Gets sent to bed" and "Gets candy," which clearly refer to consequences. The frequency of responses describing consequences rather than their preceeding behaviors seems all the more notable when consequences were not specifically solicited. It would appear, in the case of the Bad Person question especially, that for these young children it is often not a characteristic of the act itself but the consequences of the act which give the act its evaluative meaning. It is as though these children were saying, "Bad is what you get punished for" and, to a much lesser extent, "Good is what you get rewarded for." They are stating, of course, the rudiments of reinforcement learning theory. Attention is called to the significantly different frequency of responses of this sort for Good and Bad Person (4% vs. 23%, respectively). This difference may reflect a difference in the child's world between the frequency

with which good behavior is rewarded and bad behavior is punished.

The differences which occurred between Good Person and Bad Person responses in relative frequencies within the minimal/vague and plays categories also deserve comment. Responses in these categories were much more frequent in Good Person descriptions (14% and 14%, respectively) than in Bad Person responses (6% and 0%, respectively). In a sense the plays category is very close to the minimal/vague category: "Plays," "Plays nicely," and "Plays with his toys" are quite nebulous responses in terms of specific behaviors that can be identified. A marked difference occurs then in the definition of the Good and Bad Person in terms of the greater use of general, nonspecific concepts in defining Good Person.

Taken together, these results suggest that the concept of Bad Person is more focused in delineation, more specific in behavioral referents, and more frequently identified by behavior consequences than is the concept of Good Person. The concept of a good person is more vague, less specific in behavioral referents, and is less frequently identified by consequences. In terms of implications for self-concept development, the child would seem to have an easier task identifying the negative rather than the positive pole of an ideal self image and to be clearer in identifying negative rather than positive aspects of himself.

Age, Sex and Race Comparisons. Subgroup comparisons showed two significant age-related differences on the Good Person question and one such difference on the Bad Person question. Regarding Good Person definitions, prosocial responses increased with age (4% at age 3, 18%

at age 4, 21% at age 5: $\chi^2 = 6.53$, $df = 2$, $p = .04$), and responses classified as "other" decreased with age (26% at age 3, 16% at age 4, 8% at age 5: $\chi^2 = 6.76$, $df = 2$, $p = .03$). For the Bad Person question, responses in the aggressive category increased with age (35% at age 3, 39% at age 4, 58% at age 5: $\chi^2 = 8.56$, $df = 2$, $p = .02$). While these results support the developmental nature of concepts of Good and Bad Person, the absence of more age-related findings suggests that the preschool period of 3 to 6 years does not involve marked age changes and might best be characterized as a stage to be differentiated from other stages rather than a period of gradual progression. These comments are offered tentatively, for it might well be that this measure is merely insensitive to actual age progressions.

Sex and race subgroup comparisons yielded no indications of significant differences. Whatever the conceptualization of good and bad person is for these children, as measured by the present study's questions, it is common to black and white, male and female.

Summary and Recommendations

The children's responses to the Good Person and Bad Person questions indicate interesting similarities and differences. Both questions were answered predominantly in the positive mode. Defining a good person was a more difficult task: fewer children were able to give a scorable response to Good Person than to Bad Person, although the total number of responses given by children who were able to respond to the question did not differ. Good Person responses seemed less specific and consensual and most frequently referred to obedient, socially conforming

behavior. Bad Person, on the other hand, was most frequently defined in terms of aggressive behavior and as a function of consequences experienced.

The results suggest the hypothesis that defining the negative role of ideal self-image and defining negative aspects of himself may be an easier task for young children than the positive counterparts of these self-related behaviors. It is felt that these measures open up interesting further questions, particularly that of the contribution of structured group experience on these developing concepts. It is recommended that, despite their relative difficulty for children, these measures be maintained in future work. It is also suggested that the use of strictly parallel forms of the questions be investigated (i.e., good child-bad child).

CHAPTER 14

BODY PARTS

In the Body Parts component the children are asked to point to various body parts in response to the psychometrist's request. Instructions to the psychometrist are:

> Have the child get up from his chair and stand several feet from you. Stand up yourself if you feel like it. Then say: "Let's see you touch your _____." Repeat this phrase for each item. Score responses on record form.

The six items of this component are shoulder, chest, eyelash, wrist, knee, and chin. Each response is scored pass or fail by the psychometrist. Strict scoring is called for: e.g., if a child puts his hand on his thigh for "knee," it is scored as a failure. Special care is required for evaluating the response to eyelash to assure differentiation from eyelid or eyebrow responses. The total score on this component represents the number of items the child correctly identifies.

Administration and Scoring

The psychometrists felt that this component was a successful one in terms of the child's interest and cooperation. Most children liked this item and the bodily activity which was required contributed to variety within the Inventory.

A suggestion was made that the component might be improved by presenting a very easy item first and last. ("Show me your head" or "Show me your hands," for example.) The child's feeling of success could thus be enhanced.

One item, chest, was identified as a possible source of unreliability in scoring. Criteria for pass-fail on chest had not been identified prior to field work, there having been no difficulties in pilot testing. The line of demarcation between chest and stomach was a source of difficulty in several instances in the field, however, and different examiners may possibly have scored identical responses differently.

Findings

Distribution of Total Scores

The Body Parts component was administered to all 282 children. The total scores ranged from 0 to 6. The distribution of the total scores is presented in Figure 20. The distribution was somewhat skewed in the direction of higher scores. While skewness of any sort is not desirable psychometrically, the rapport that occurs when a component is not "hard" or discouraging for most of the subjects compensates somewhat. The mean of the total scores was 3.9, with a standard deviation of 1.4.

Item Analyses

Considered here and presented in Table 48 are results which allow the internal consistency of this component to be inspected: the difficulty of each item and its discriminability. Item difficulty is represented by the percentage of children passing the item. Item discriminability is represented by a biserial correlation and is reflected by the average total scores of children who pass the item and those who fail the item. It can be seen in Table 48 that item difficulties ranged from 27%

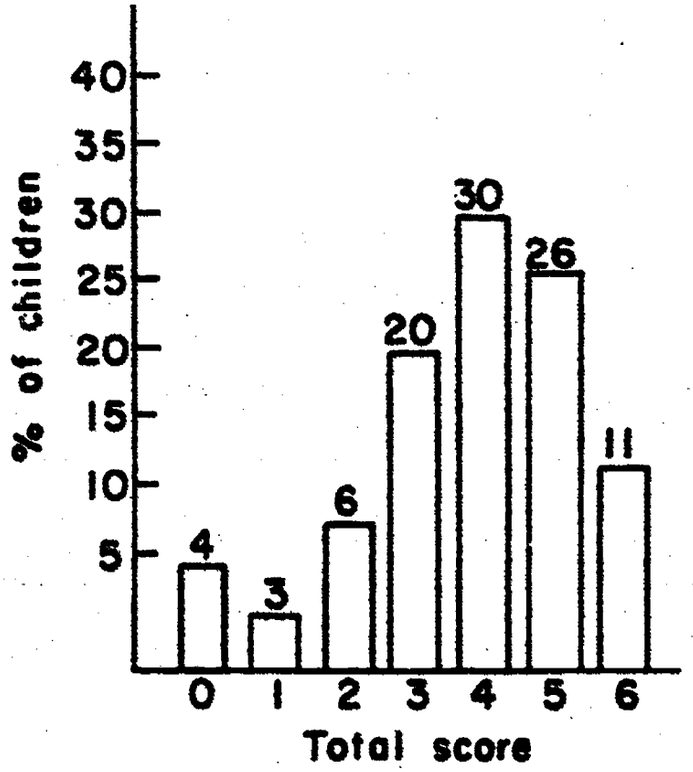


Figure 20. Distribution of the total scores of the component Body Parts.

TABLE 48

ITEM ANALYSES OF BODY PARTS

Item (in order of increasing difficulty	% of children passing item	Biserial correlation with total score	Average total score of children who pass item	Average total score of children who fail item
Knee	91	.91	4.2	1.2
Shoulder	84	.89	4.3	1.9
Chin	84	.88	4.3	1.9
Eyelash	57	.77	4.7	2.9
Chest	49	.66	4.7	3.2
Wrist	27	.71	5.2	3.4

passing to 91% passing. This spread of difficulty in the items could be improved (a) by adding an item in the very difficult category (<20%) and by adding items to the moderate range (60% to 70%) and (b) by removing one of the two items which had a difficulty level of 84%. The item correlations with total score were very satisfactory. They were all positive and quite high. While higher correlations might possibly be hoped for, the restriction in the range of possible scores for both the items and the total score impedes the likelihood that improvement could be substantial. The discriminability of each item is further illustrated by the difference in total scores of children who passed the item and of children who failed the item. These differences were significant beyond the .01 level for each item.

The interitem consistency of this component, as assessed by the Kuder-Richardson formula 20, was .39, indicating that there was not a great deal of homogeneity in the items.

Subgroup Analyses of Total Scores

Table 49 presents the means and standard deviations of the Body Parts total score for the total sample and for age, sex, and race subgroups together with the significance and (p) for tests of subgroup differences.

Inasmuch as a performance on Body Parts is assumed to reflect cognitive differentiation of physique elements, the total score should increase with age. The statistically significant differences among the age groups, then, leads to confidence in the measure. An analysis of variance indicated that significant differences appeared in the age subgroups ($F = 11.47$, $df = 2/279$, $p < .01$). Tukey's test

TABLE 49

SUBGROUP MEANS AND STANDARD DEVIATIONS FOR

THE TOTAL SCORES ON BODY PARTS

Group	N or n	$\bar{X}$	SD	p
All children	282	3.9	1.4	
Age: 3	66	3.3	1.5	
4	142	3.9	1.4	.01
5	74	4.6	1.1	
Sex: Male	143	3.9	1.4	.98
Female	139	3.9	1.5	
Race: White	73	4.2	1.1	.05
Black	86	3.8	1.4	

subsequent to this analysis indicated that significant differences existed between each age subgroup mean ($3 < 4 < 5$).

Predictions regarding sex and race subgroup differences in total score were not made. No sex differences were found. Differences appeared, however, in the race comparison. White children scored higher than black children ($t = 2.02$, $df = 156$, $p = .05$). Interpretation of this finding awaits further analyses.

Recommendations

It is recommended that this component be retained in the Inventory with the following revisions:

a. For the sake of rapport this component should begin and end with a very easy item which all subjects can pass. These items need not enter into the scoring of the component.

b. Items should be added and/or substituted in order to increase the range in difficulty, and the representation of the middle range of difficulty should be increased.

c. A definite pass-fail criterion for scoring "chest" should be set; the area between waist and neck would seem a clear criterion.

CHAPTER 15

WISH

The Wish question was included in the Inventory to explore expressions of personal desires, one aspect of the child's phenomenological self. The text of the Wish question is as follows:

> Let's pretend you can wish for anything you want...anything in the whole world. If you had a wish and your wish could come true, what would you wish for?

Following the child's response, the examiner asks for the reason(s) for the wish ("Why would you want _____?") and, if necessary, clarifies the identity of the thing wished for.

Administration and Scoring

The examiners reported that this question seemed quite difficult for the children. Examiners sensed that most children seemed unable to adopt the reflectivity and long-range orientation which would characterize the response of an older child or adult. Unable to comprehend or cope with selecting "anything in the whole world" some children reportedly scanned the room for situational cues to a response. One examiner reported she felt this question "separated" the children into those who could handle the reflectivity required and those who were nonplussed by it.

This question was felt to be a poor choice for the last component of the Inventory, particularly for those children who had difficulty with it. Any series of measures utilized in the future should conclude with a "success item" for all children.

The children's wishes and reasons were categorized according to the content observed in the responses. The reliability of such scoring was high. Two scorers independently scoring 20 records achieved 95% agreement.

Findings

The findings regarding the number of children able to give a scorable response to the Wish question support the examiners' impressions of the relative difficulty of this question. Of all children asked the question, only 78.9% were able to give a scorable reply. Ability to answer the question was not related to sex or race but was significantly related to age ($\chi^2 = 13.50$, $df = 2$, $p = .001$). At 3 years of age 67.2% of the children could answer the question, at 4 years of age, 77.0%, and at 5 years of age, 92.9%.

Children giving scorable responses mentioned an average of 1.2 items they would wish for. No sex or race differences appeared in this figure, and age differences, while significant ($F = 3.88$, $df = 2/207$, $p < .025$), did not point to large differences nor consistent age trends. Four-year-olds gave more responses (1.3) than either 3- or 5-year-olds (1.1 each).

Table 50 presents the categorized responses of the children to the Wish question. The categories have been grouped to parallel classifications used by others in investigating the expressed wishes of older children. Included is the notation of any age, sex, or race differences which were found in the frequency of any response.

It can be seen from Table 50 that material objects or possessions were the most frequent wishes mentioned and that food and toys lead

TABLE 50

RESPONSES TO THE WISH QUESTION

Response	Percentage of children mentioning[a]	Subgroup differences
Material objects or possessions		
Food	16.2	
Toys: sex stereotyped (e.g., dolls, for girls, toy trucks for boys)	12.0	
Toys: sex neutral (e.g., bicycle, swing)	11.0	Sex: girls > boys., $p = .04$
Adult supplies (e.g., camera telescope)	7.1	
Vehicles	5.7	Sex: boys > girls, $p = .0007$
Money	2.9	Age: $3 < 4 < 5$, $p = .02$
Clothing	2.9	
Housing and household goods	2.9	
Self-enhancement or development		
Learn (e.g., to read, to draw)	4.3	
Be or be like fantasy character (e.g., Batman)	3.3	
Improved stature or physical skill	1.9	
Relationships and companionship		
Related to peer	3.8	
Related to father	1.4	
Related to mother	1.0	
Other		
Pet	15.2	
Pleasant event (e.g., birthday)	5.2	
Fantasy character (e.g., wish for the Easter Bunny)	4.3	
Other	12.0	

[a]Percentage total exceeds 100 because of multiple responses.

the list of things desired. Wish for a pet was also frequently mentioned; wishes related to self-enhancement and development and to relationships with others were relatively infrequent. Attention is also drawn to the fact that subgroup differences were scent: the items wished for did not vary appreciably by age, sex, or race.

Interpretation of these data is hazardous without comparative data. Unfortunately, no responses from children without group care experience are available, and previous similar research has involved only older children. This previous research has tended to show, as in this sample, a preponderance of wishes for material possessions (which decline with age) and more desires in the self and relationship categories (which increase with age). It may be that our data represent simply a downward extension of trends found with older children. Most parsimoniously they should be regarded as such. At some point, however, it may be advisable to investigate if the prevalence of creature comfort items, food in particular, is merely a reflection of the developmental stage of the children or reflects a dimension of deprivation that would not appear in other groups of children of the same age. It would be interesting also to investigate under what conditions wishes regarding the self are increased or diminished.

The reasons children gave for their wishes referred primarily to the appropriate use of items (e.g., "to ride it") or to global affect associated with the item (e.g., "because I like dogs"); thus the content of these reasons did not prove particularly enlightening or interesting. There were no subgroup differences in reasons given.

Recommendations

The Wish question was found to be a difficult one for preschool children. The findings suggest, however, that it may be useful to retain this question if, in future work, group care and nongroup care children are to be compared. It should not be the concluding component in the Inventory; the final component should be an enjoyable item on which all children can easily feel they have performed successfully.

CHAPTER 16

CAMERA GAME

Camera Game attempts to assess expressive aspects of affective development in children. Additionally, it provides an opportunity to judge the child's role playing ability and his enjoyment of fantasy or pretending. In Camera Game, the child is asked to reproduce five mood expressions (happy, sad, sleepy, mad and scared) while the psychometrist pretends to take his picture. The procedure and instructions are as follows:

> Say, "Now let's play another pretend game. Let's pretend that I have a camera and that I'm going to take some pictures of you. Here is my camera." (Pretend that you are getting ready to take picture of child. Hold up your hands to form a camera and aim it at the child. Say, "Click," and move your index finger as if pushing the shutter after he has produced each response. It is important that you enter into this game with gusto and make it fun for the child and for yourself.)

After demonstrating the pretend camera, examiners are told to introduce the first mood stimulus as follows:

> First, I want to take a picture of you when you're very happy. Show me what you look like when you're very happy!

After the subject produces his response, examiners are instructed to say, "Click," in order to emphasize the "let's pretend" atmosphere. Instructions for other mood stimuli parallel those for "happy."

Responses are scored by the psychometrist in the following categories:

+ A really good response. Easily recognizable without knowledge of stimulus word.

± A marginal response. Fairly good but some ambiguity possible without knowledge of stimulus word.

- A response is produced but, in terms of the stimulus word, is unrecognizable or inappropriate.

NR No response.

An additional scoring category is utilized for noting gross motor components when a child utilizes postural and gestural means of expression. Finally, the child is rated on his enjoyment of the pretending involved (from 1, seems to relish pretending, to 3, seems uninvolved with fantasy aspects of game).

Administration

The examiners' feedback regarding Camera Game indicated some vagaries in introducing the child into the task. Some children reportedly wanted to do the pretend picture taking, and some gave verbal associations rather than affective display (e.g., "Mommy got mad at me" for the stimulus mad). Such difficulties, however, were not extensive.

The examiners stated they felt they were consistent over time and children in judging expressions, gross motor involvement, and fantasy enjoyment. Some wondered, however, if they would be consistent with other examiners in their scoring. The issue, of course, is that of interscorer reliability, on which pilot data had been obtained prior to utilizing Camera Game in the Inventory. These data provided confidence that different raters made sufficiently similar scores. The concern expressed by the examiners emphasizes the fact that particular attention should be given this matter in further development of Camera Game and in future training sessions with psychometrists who will administer the Inventory.

Findings

Distribution of Total Scores

The Camera Game was administered to 272 children. The remaining 10 children in the total sample of 282 are those with whom administration of the Inventory was discontinued prior to the appearance of Camera Game in the sequence of components (third from last). The total scores ranged from 0 to 5. Figure 21 presents the distribution of these total scores. It can be seen that the distribution by no means approached the desired normal distribution. There was too high a percentage of children with ceiling and floor scores (all right or all wrong), and the peak comes at one extreme rather than centrally.

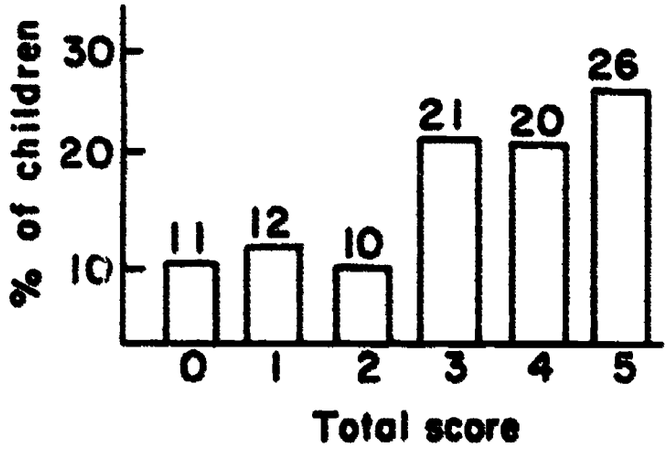


Figure 21. Distribution of the total scores of the Camera Game component.

Item Analyses

Table 51 presents the item analysis data for Camera Game. Item difficulties, represented by the percentage of children succeeding on each item, ranged from 45% to 78%. This range of difficulty is too restricted and too skewed in the high direction. This component should thus be improved by adding items to achieve a greater range of difficulty and by adding one item which centers around the 50% difficulty level.

TABLE 51

ITEM ANALYSES OF CAMERA GAME

Item (in order of increasing difficulty)	% of children passing item	Biserial correlation with the total score	Average total score of children who pass the item	Average total score of children who fail the item
Scared	45	.87	4.3	2.0
Sad	53	.84	4.1	1.9
Mad	57	.95	4.2	1.6
Sleepy	74	.99	3.8	1.0
Happy	78	.87	3.6	1.0

The discriminability of each item, reflected in its high correlation with the total score is excellent. Comparison of total scores of those who passed the item and those who failed the item further illustrates this discriminability: each difference was significant at or beyond the .05 level. Thus, each item was highly successful in its contribution to the total score.

The interitem consistency of this component as assessed by the Kuder-Richardson formula 20 was .60, indicating that there was moderate homogeneity in the items.

Subgroup Analyses of Total Scores

Camera Game total scores (means and standard deviations) are given in Table 52 for the total sample and for age, sex, and race subgroups. Referring to Table 52, the skewness of the distribution which was illustrated in Figure 21 is again apparent. The mean score was 3.0 for this component, which has a maximum possible score of 5. The standard deviation of 1.7 indicates acceptable variability.

TABLE 52

SUBGROUP MEANS AND STANDARD DEVIATIONS OF

THE TOTAL SCORES ON CAMERA GAME

Group		N or n	$\bar{X}$	SD	p
All children		272	3.0	1.7	
Age:	3	64	2.2	1.7	
	4	137	3.2	1.6	.01
	5	71	3.5	1.5	
Sex:	Male	138	3.0	2.0	.99
	Female	134	3.0	1.9	
Race:	White	71	3.4	1.9	.87
	Black	82	3.4	2.0	

An increase in the mean total score with age was expected on the basis that the ability to produce emotional expressions accurately develops with age. The age subgroup data in Table 52 indicate that this expected increase was found. An analysis of variance indicated that significant differences existed in the means ($F = 11.33$, $df = 2/269$, $p < .01$). Further analyses by Tukey's test of age differences in these means indicated that the difference between the 4- and 5-year-olds, however, did not reach significance ($3 < 4 = 5$).

No sex or race differences in total score were predicted and none were found.

Summary and Recommendations

An improved Camera Game should be included in further work with the Inventory. This component reflects development with age and its individual items possess splendid discriminability. Its difficulty range and skew, however, represent serious shortcomings, and confidence in interrater reliability may profit from further attention.

The following specific suggestions are made:

a. Additional items should be developed so that range and mean difficulty level approach the optimal more closely.

b. Further interscorer reliability studies should be conducted, and any appropriate revisions which would increase this reliability should be made in the procedure.

c. Special efforts should be made to assure acceptable interrater reliability of field psychometrists, and their confidence in their reliability, during training sessions, i.e., train until a satisfactory level is achieved.

APPENDIX

GLOSSARY OF STATISTICAL TERMS AND SYMBOLS

Glossary of Statistical Terms and Symbols

Term and/or symbol	Definition
correlation (r)	A correlation is a number expressing to what extent two sets of scores are related or vary together. The possible range of a correlation coefficient (r) is from 1.00 (a perfect positive relationship between two sets of scores) through 0 (no relationship) to -1.00 (a perfect negative relationship). The numerically larger the correlation, the greater the relationship.
df	Degrees of freedom. This number is used in determining the probability (p) of a statistical finding (e.g., a t, F, or χ^2). This term is included in this report for technical accuracy and can be disregarded by the nontechnical reader.
F	A statistic used to judge whether there are statistically significant differences in scores among three or more groups. The F and its df, reported for the technical reader, can be disregarded by the nontechnical reader. They are used to determine the probability (p) that the findings can be regarded as statistically significant or merely due to chance.
McNemar's test	This test is used to compare the proportion of responses in categories which are common to two related questions. The test is a form of the χ^2 test.
Mean ($\bar{X}$)	The sum of scores or amounts divided by the number of scores or amounts. The nontechnical term "average" most commonly refers to the mean, although there are other types of averages.
N, n	Number of individuals (e.g., children, parents) in a sample. As used in this volume, N refers to a large sample (e.g., all children who answered a particular question) and n refers to a subsample within the larger sample (e.g., boys or girls).

Term and/or symbol	Definition
ns	Statistically nonsignificant
p	Probability or significance level. This number expresses the likelihood that the results could have occurred "by chance." It is expressed as a decimal fraction. A p of .05, for example, means that the results are likely to be achieved "by chance" only 5 times out of 100; a p of .001 means that the results could have arisen "by chance" only one time out of 1000. A p of .05 or less is generally taken to indicate statistically significant findings -- those which one feels fairly confident in assuming did not arise by chance. In this volume the convention of referring to p's between .05 and .15 as trends toward statistical significance has been adopted. Probabilities (p's) greater than this are regarded as statistically nonsignificant.
range	The distance between the lowest score and the highest score.
SD	Standard deviation. This term is an index of the dispersion or spread-outness of a set of scores. For example, two sets of scores might have a mean score of 10. One set might have an SD of 2 and the other set might have an SD of 3; the second set has a broader dispersion of scores. Roughly 2/3 of the scores occur in the range of +1 SD above the mean to -1 SD below the mean.
t	A statistic relating to the t-test which is used to test whether there are statistically significant differences in scores between two groups. The t and its df (which can be disregarded by the nontechnical reader) are used to determine the probability (p) that the differences in scores are significant or merely due to chance.

Term and/or symbol	Definition
χ^2	Chi-square. The chi-square test is applicable when there are response categories rather than scores and one wishes to determine if two or more groups differ significantly with regard to their response patterns (e.g., Does one group choose some categories more often than another group?). The χ^2 and its df, which can be disregarded by the nontechnical reader, are used to determine the probability that the findings are due merely to chance or can be regarded as "significant."
<	Less than. For example, < 10 is read "less than 10."
>	Greater than. For example, > 10 is read "greater than 10."

References

American Psychological Association. Technical recommendations for psychological tests and diagnostic techniques. Washington: Author, 1954.

Anastasi, A. Psychological testing. (3rd ed.) New York: MacMillan, 1968.

Insell, E. M., & Hansen, J. C. Patterns in vocational development of urban youth. Journal of Counseling Psychology, 1971, 18, 505-508.

Banta, T. J. Tests for the evaluation of early childhood education: The Cincinnati autonomy test battery. In J. Hellmuth (Ed.), Cognitive studies. Vol. I. New York: Bunner/Mazel, 1970.

Dubin, R., & Dubin, E. Children's social perceptions: A review of research. Child Development, 1965, 36, 809-838.

Erikson, E. H. Childhood and society. (2nd ed.) New York: Norton, 1963.

Finch, H. Young children's concepts of parental roles. Journal of Home Economics, 1955, 47, 99-103.

Gish, D. J. An evaluation model for early childhood education. Unpublished doctoral dissertation, The Pennsylvania State University, 1971.

Hartley, R. Children's concepts of male and female roles. Merrill-Palmer Quarterly, 1960, 6, 83-91.

Hauser, S. T. Black and white identity formation: Studies in the psychological development of lower socioeconomic class adolescent boys. New York: Wiley-Interscience, 1971.

Havighurst, R. J. Youth in exploration and man emergent. In H. Borow (Ed.), Man in a world at work. Boston: Houghton Mifflin, 1964.

Looft, W. R. Sex differences in the expression of vocational aspirations by elementary school children. Developmental Psychology, 1971, 5, 366. (a)

Looft, W. R. Vocational aspirations of second-grade girls. Psychological Reports, 1971, 28, 241-242. (b)

Mussen, P. H. (Ed.) Carmichael's manual of child psychology. (3rd ed.) New York: John Wiley & Sons, 1970.

Peters, D. L., Hendrickson, J., Marcus, R., & Ridley, P. Defining day care goals: A preliminary study. (The Pennsylvania Day Care Study Technical Report No. 5) University Park, Pa.: The Pennsylvania State University, College of Human Development, Institute for the Study of Human Development, 1972.

Peters, D. L., & Stein, N. Project head start--Summer 1966: An evaluational report. San Mateo, Calif.: San Mateo Human Resources Commission, 1966.

Schaefer, E. S., & Aaronson, M. Day care behavior inventory. Unpublished rating scale. Rockville, Md.: National Institute of Mental Health, Department of Health, Education and ' care, no date.